

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Economics

2016

Essays in Health Economics

Alexander, Diane

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp011544br515>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

ESSAYS IN HEALTH ECONOMICS

Diane Elizabeth Alexander

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS
ADVISER: Janet M. Currie

May 2016

© Copyright by Diane Elizabeth Alexander, 2016. All rights reserved.

Abstract

Health is a key determinant of the quality of life. As such, the extremely high cost of health care in the US and the resulting barriers to access are concerning. This dissertation consists of three essays exploring the determinants of health care provision—both in terms of access and type of treatment—and their impact on patient health and health care costs.

In Chapter 1, I investigate how policies that pay physicians to reduce costs affect treatment decisions. I show that a Medicare pilot program in New Jersey that paid doctors bonuses for reducing costs was unsuccessful; doctors changed which patients were admitted and sorted healthier patients into participating hospitals, but did not reduce costs. I argue that understanding the gaming behavior displayed by physicians is important for policy design.

While Chapter 1 explores how doctors make decisions over treatment, the second chapter shifts focus to the determinants of access to treatment itself. In Chapter 2, co-authored with Molly Schnell, we link the supply of available providers to the provision of mental health care and mental health outcomes. While the costs of untreated mental illness are widely recognized, access to treatment is often limited. In this chapter, we show that allowing nurse practitioners to prescribe medication without physician oversight is associated with improvements in population mental health. Our results suggest that liberalizing scope of practice legislation for nurse practitioners can extend care to both underserved and disadvantaged populations.

In the third chapter, I turn back to physicians, and examine how payment can affect both treatment choice and patient health. In particular, I explore how changing pay affects the decision of whether or not to perform a C-section. I find that when Medicaid pays physicians relatively more for C-sections, physicians perform more C-sections, and there are fewer infant deaths. Taken together, these findings suggest that C-section rates may be too low for some groups, and that policies aimed at decreasing costs by lowering procedure use among low-income patients may have adverse health consequences.

Acknowledgements

I am deeply indebted to my advisor, Janet Currie, whose selfless gifts of time and seemingly limitless support kept me motivated over the past six years. I am grateful to Ilyana Kuziemko, whose return to Princeton was extremely fortuitous—her excellent advice and attention to detail greatly improved the quality of my research. I also thank Hank Farber, who has known me since before I was a student, for his support, wisdom, and advice throughout my entire time at Princeton. I thank my perpetual coauthor, Molly Schnell, for being a fantastic researcher and sounding board for new ideas, as well as a wonderful friend.

I am forever grateful to participants in the Center for Health and Wellbeing and the Research Program in Development Studies Weekly Seminar Series, the Public Finance Working Group, the Joint degree Program in Social Policy, and the Fellowship of Woodrow Wilson Scholars at Princeton University for conversations and suggestions that have greatly improved the quality of this project. I would especially like to acknowledge Anne Case, Angus Deaton, Will Dobbie, Thomas Fujiwara, Jean Grossman, Jon Gruber, Fernanda Marquez, Alex Mas, Christopher Neilson, Uwe Reinhardt, Harvey Rosen, David Slusky, Tom Vogl, and Dan Zeltzer.

I am thankful for financial support from the Center for Health and Wellbeing Program on U.S. Health Policy, the Fellowship of Woodrow Wilson Scholars, the Griswold Center for Economic Policy Studies, and the Joint Degree Program in Social Policy at Princeton University, as well as the The Dwight D. Eisenhower/Clifford Roberts Graduate Fellowship Program. I also thank the staff at the New Jersey Department of Health for their help in accessing data.

I would not have made it through the past six years without the companionship of my classmates, who made graduate school one of the most rewarding experiences of my life. I would particularly like to thank Pauline Leung, Inessa Liskovich, Moshe Katzwer, and Sharon Traiberman: their support, advice, and friendship kept me sane and grounded (or as much as can be expected). I thank Tom Winberry for the steady supply of optimism, encouragement, and love that kept me afloat during the hard times, and made the good times even better. Most of all, I am thankful for the love and support of my parents, Helen and Dave Alexander.

Contents

Abstract	iii
Acknowledgements	iv
1 How do Doctors Respond to Incentives? Unintended Consequences of Paying Doctors to Reduce Costs	1
1.1 Introduction	1
1.2 Institutional Background	6
1.2.1 How Doctors Treat Patients within Hospitals	6
1.2.2 How Doctors and Hospitals are Paid	7
1.2.3 Implementation of Gainsharing Demonstration	10
1.3 Conceptual Framework	11
1.3.1 The Set Up	12
1.3.2 No Bonuses	13
1.3.3 With Bonuses	15
1.4 Data and Empirical Strategy	18
1.4.1 Data Sources	19
1.4.2 Estimation Sample	20
1.4.3 Empirical Strategy	20
1.5 Results	22
1.5.1 The Admission Margin	22
1.5.2 The Hospital Sorting Margin	24

1.5.3	The Quantity of Services Margin	28
1.5.4	Extensions and Robustness Exercises	30
1.6	Conclusion	34
1.7	Figures	36
1.8	Tables	42
1.9	Appendix	51
1.9.1	Appendix Figures	61
1.9.2	Appendix Tables	66

2 Just What the Nurse Practitioner Ordered: Independent Prescriptive Authority and

Population Mental Health	77
2.1	Introduction 77
2.2	Data 81
2.2.1	Independent Prescriptive Authority 81
2.2.2	Mental Health Outcomes 82
2.2.3	Health Resources 85
2.2.4	Prescription Data 86
2.3	Prescriptive Authority and Mental Health Outcomes 87
2.3.1	Mental-Health-Related Mortality 87
2.3.2	Self-Reported Mental Health 90
2.4	Prescriptive Authority and Prescribing Patterns 92
2.5	Conclusion 94
2.6	Figures 96
2.7	Tables 97
2.8	Appendix 101
2.8.1	Appendix Figures 102
2.8.2	Appendix Tables 103

3 Does Physician Pay Affect Procedure Choice and Patient Health? Evidence from Medicaid C-section Use	118
3.1 Introduction	118
3.2 Conceptual Framework	121
3.3 Data	123
3.3.1 Vital Statistics Data	123
3.3.2 Physician Reimbursement Data	125
3.3.3 Hospital Data	127
3.4 When Should a C-section Be Performed?	128
3.5 Empirical Strategy	130
3.6 Results	134
3.6.1 Impact of Pay Difference on Procedure Choice	134
3.6.2 Impact of Pay Difference on Infant Health	137
3.6.3 Pathways between pay differential and infant health	139
3.7 Conclusion	141
3.8 Figures	143
3.9 Tables	147
3.10 Appendix	155
3.10.1 Appendix Figures	157
3.10.2 Appendix Tables	161
Bibliography	172

List of Tables

1.1	All Patient Refined-Diagnosis Related Group (APR-DRG) Examples	42
1.2	Hospital Characteristics	43
1.3	Main Sample Characteristics	44
1.4	Bonuses Change Which Patients are Admitted	45
1.5	Doctors Sort Healthier Patients into Participating Hospitals	46
1.6	Doctors Sort Healthier Patients into Participating Hospitals: High Admission Di- agnoses	47
1.7	Effect of Program on Population-Level Severity of Illness	48
1.8	Bonuses Do Not Reduce Costs or Change Procedure Use	49
1.9	Replicating the Naive Evaluation: Medical Patients	50
1.10	Simulated Share Treated on Costs and Procedure Use	50
1.11	Top 15 Diagnoses for Medical and Surgical Patients	66
1.12	Participating Hospitals	67
1.13	Effect of Program on Ex-Ante Patient Health: Medicare ER Patients	68
1.14	Effect of Program on Baseline Admission Probability: Medicare ER Patients . . .	69
1.15	Effect of Program on Costs and Procedure Use: Medicare ER Patients	70
1.16	Effect of Program on Baseline Admission Probability: Near Medicare Patient . . .	71
1.17	Effect of Program on Ex-Ante Patient Health: Near Medicare Patients	72
1.18	Effect of Program Costs and Procedure Use: Near Medicare Patients	73
1.19	Effect of Program on Baseline Admission Probability: Single Hospital Doctors . .	74

1.20	Effect of Program on Ex-Ante Patient Health: Single Hospital Doctors	75
1.21	Effect of Program Costs and Procedure Use: Single Hospital Doctors	76
2.1	Summary Statistics: County-Level Mortality and Controls, 1990 - 2014	97
2.2	Summary Statistics: Self-Reported Mental Health and Controls, 1993 - 2013	98
2.3	Independent Prescriptive Authority and Mental-Health-Related Mortality	99
2.4	Independent Prescriptive Authority and Self-Reported Mental Health	100
2.5	Independent Prescriptive Authority and Prescriptions by MDs	100
2.6	Robustness: Correlates of Law Changes	103
2.7	Missing BRFSS State-Years	104
2.8	Summary Statistics: Self-Reported Mental Health and Controls, 1993 - 2013	105
2.9	Summary Statistics: County-Level Mortality and Controls, 1990 - 2014	106
2.10	Prescriptive Authority and Suicides	107
2.11	Prescriptive Authority and Mental-Health-Related Mortality	107
2.12	Mental-Health-Related Mortality: Full Regression Results	108
2.13	Prescriptive Authority and Mortality: Unweighted	109
2.14	Suicides Robustness (Leave Out)	110
2.15	Mental-Health-Related Mortality Robustness (Leave Out)	111
2.16	Mental-Health-Related Mortality: Including State-Specific Linear Time Trends . .	112
2.17	Prescriptive Authority and Days in Poor Mental Health	112
2.18	Prescriptive Authority and 21+ Days in Poor Mental Health	113
2.19	Self-Reported Mental Health: Full Regression Results	114
2.20	Self-Reported Number of Days in Poor Mental Health: Robustness (Leave Out) . .	115
2.21	Self-Reported 21+ Days in Poor Mental Health: Robustness (Leave Out)	116
2.22	Self-Reported Mental Health: Including State-Specific Linear Time Trends	117
3.1	Payment Variable Means	147
3.2	Summary Statistics	148

3.3	Probability of Primary C-section by Insurer in 2012 (x100)	149
3.4	Effect of Salary Structure and Pay Differentials on C-Section Use (x100)	150
3.5	Effect of Salary Structure and Pay Differentials on Infant Death (x100)	150
3.6	Effect of Salary Structure and Pay Differentials on Neonatal Death (x100)	151
3.7	Effect of Salary Structure and Pay Differentials on Death After Neonatal Period (x100)	151
3.8	Effect of Pay Difference on Infant Death (x100)	152
3.9	Effect of Pay Difference on Neonatal Death (x100)	152
3.10	Effect of Pay Difference on Death After Neonatal Period (x100)	153
3.11	Effect of Pay Difference on Birth Weight	153
3.12	Effect of Pay Difference on Gestational Age at Birth	154
3.13	Effect of Pay Difference on Number of Prenatal Visits	154
3.14	Medicaid Fee Schedule Data Used	161
3.15	Components of Salary for Military and IHS Doctors	162
3.16	Women with a Previous C-section: Effect of Salary Structure and Pay Differentials on C-Section Use (x100)	163
3.17	Women with a Previous C-section: Effect of Pay Differentials on Infant Death (x100)	163
3.18	Stratifying by Medicaid Managed Care: Effect of Salary Structure and Pay Differ- entials on C-Section Use (x100)	164
3.19	Effect of Salary Structure and Pay Differentials on C-Section Use: Linear State Trends (x100)	165
3.20	Effect of Salary Structure and Pay Differentials on Infant Death: Linear State Trends (x100)	166
3.21	Procedure and Billing Descriptions	167
3.22	Relationship between Pay Differentials and Unemployment	168

List of Figures

1.1	Hypothetical Bonus Calculation	36
1.2	Distribution of Maximum Bonuses	36
1.3	Hospital Locations	37
1.4	Doctor's Utility as a Function of β : without Bonuses	38
1.5	Doctor's Utility as a Function of β : with Bonuses	39
1.6	Optimal Quantity of Care as a Function of β	40
1.7	Healthier Patients Sent to Participating Hospitals	41
1.8	Doctor's Utility as a Function of β : without Bonuses	61
1.9	Doctor's Utility as a Function of β : with Bonuses	62
1.10	Optimal Quantity of Care as a Function of β : with Bonuses	63
1.11	Placebo Test: Randomize Program Participation	64
1.12	Placebo Test: Randomize Dates	65
2.1	Changes in Independent Prescriptive Authority for NPs: 1990 – 2014	96
2.2	Percent of US Population Living in Underserved Counties: 1990 – 2014	96
2.3	Mental Health Services by County: 2002	102
2.4	Mental Health Services by State: 2002	102
3.1	Medicaid Coverage of Births in the US	143
3.2	Variation in Pay Difference over Time	144
3.3	Locations of Federal Government Controlled Hospitals	145

3.4	Case Study: Arkansas	146
3.5	Estimates from Randomly Assigning States to Pay Schedules	157
3.6	Variation in Pay Difference over Time	158
3.7	Case Study: Arkansas (Infant Death)	159

Chapter 1

How do Doctors Respond to Incentives? Unintended Consequences of Paying Doctors to Reduce Costs*

1.1 Introduction

Lowering the growth in health care costs has long been a key U.S. public policy goal. Yet while many ideas exist for how to reduce costs, there is no consensus on which path is most promising Gruber (2008, 2010). Because of this uncertainty, the Patient Protection and Affordable Care Act (ACA) has earmarked billions of dollars for pilot programs.¹ The ACA's strategy is to try “virtually every cost-control reform proposed by doctors, economists, and health policy experts and [include] the means for these reforms to be assessed quickly and scaled up if they're successful,” thus ensuring “that effective change will occur” Orszag and Emanuel (2010). A large set of these pilot programs focus on changing the financial incentives of doctors, motivated by the fact that

*Materials from this chapter were presented at the Annual Meetings of the Society of Labor Economists (2015), the Annual Meetings of the American Economics Association (2015), and the Annual Meetings of the Association for Public Policy Analysis and Management (2014). I thank Anne Case, Janet Currie, Hank Farber, Jon Gruber, Ilyana Kuziemko, Molly Schnell, Sharon Traiberman, Tom Vogl, Thomas Winberry, and participants in the Center for Health and Wellbeing and Public Finance workshops at Princeton University for helpful comments. I also thank the staff at the New Jersey Department of Health for their help accessing the data. This research received financial support from the Center for Health and Wellbeing.

¹The Center for Medicare and Medicaid Innovation was established by Section 3021 of the Affordable Care Act (ACA). The Innovation Center is tasked with testing innovative health care payment and service delivery models with the potential to improve the quality of care and reduce Medicare, Medicaid, and CHIP expenditures. The ACA appropriated \$10 billion for the Innovation Center from FY 2011 to FY 2019 (<http://www.hhs.gov/about/budget/fy2015/budget-in-brief/cms/innovation-programs/index.html>).

doctors in the US are paid separately for each service provided (“fee-for-service”), potentially encouraging them to perform unnecessary procedures. These programs offer the opportunity to study how much and on which margins doctors respond to altered payment schemes, an important open question in the literature. However, the small-scale nature of pilot programs leaves them susceptible to gaming and selection bias, making it unclear whether the information they generate is actually informative for a nationwide reform.

In this paper, I study a pilot program that paid doctors for reducing costs. I ask both how doctors responded to these changing incentives, and whether the information learned is useful for informing a national reform. In particular, I analyze the effects of the New Jersey Gainsharing Demonstration, under which hospitals paid doctors bonuses for reducing the total costs of treatment for each Medicare admission. The bonuses were tied to the total costs incurred during a hospital stay, and were designed to experiment with bundled payments—an incentive scheme where doctors are paid one fee for treating a patient, rather than separately for each service provided.

I find three main results of this program. First, doctors change which patients are admitted, as opposed to treated and then sent home. In particular, doctors are more likely to admit patients whose treatment generates high expected bonuses, and less likely to admit patients who generate low expected bonuses. Second, doctors often work in more than one hospital, and thus can change where patients are admitted. Doctors sort healthier patients into participating hospitals (even conditional on the program’s risk-adjustment criteria), as these patients are cheaper to treat and therefore generate bigger bonuses. Third, conditional on admission and patient health, doctors do not change their procedure choice or otherwise lower treatment costs. Thus, the bonuses caused doctors to change their admission decisions—where and whether a patient was admitted—rather than reduce costs.

My empirical strategy leverages the fact that many doctors treat patients in more than one hospital. I measure the effect of the bonuses by comparing changes in a doctor’s behavior at a participating hospital to the same doctor in a non-participating hospital. I worked at the New Jersey Department of Health to construct a unique dataset which allows me use a within-doctor

specification. These data allow me to follow both both patients and doctors over time, and across all hospitals in New Jersey. The data include admission and discharge dates, all diagnoses and procedure codes, payer and patient demographic information, codes for doctors and patients, and list charges.

The bonuses given to doctors under the pilot program were designed to lower costs by reducing the incentive to provide treatments with low marginal benefits. In practice, patients were divided into types by diagnosis and severity of illness, and a maximum bonus was assigned to each group. Doctors were then paid a fraction of this maximum bonus after treating an eligible patient, depending on how close they got to pre-program cost benchmarks. Because of the limited scope of the pilot program, doctors could only receive a bonus if they treated an admitted Medicare patient at a participating hospital.

My first main result is that the cost-reduction bonuses change the patterns of admission across patient types. Admission is an important outcome, both in terms of costs and patient health; it is the difference between intense and prolonged monitoring by health care professionals, and being sent home after treatment.² I find that the cost reduction bonuses are associated with an increase the admission of patients in high-bonus types, relative to baseline. Conversely, patients in low-bonus types are less likely to be admitted. As capacity constraints and program rules limit the ability of doctors to increase overall admission rates, doctors instead reallocate admission across patients.

Second, doctors send healthier patients to participating hospitals in response to the program. After the bonus program was implemented, the mix of patients who were admitted at participating hospitals were ex ante healthier. Patients admitted to participating hospitals had fewer chronic conditions and lower scores on co-morbidity indices, conditional on their type. As healthier patients are cheaper to treat, doctors receive higher average bonuses for treating these patients. While defining the bonuses within diagnosis and severity level cells was meant to serve as a type of risk-adjustment, doctors were able to identify low-cost patients even within these groups, and sort patients across hospitals in order to increase their expected bonus payments.

²Generally, admitted patients are assigned a hospital bed and spend at least one night at the hospital.

Finally, conditional on admission and patient health, the bonuses did not reduce costs or change procedure use. I look at many measures of services performed: length of stay, the use of different types of diagnostic imaging procedures labeled as overused by doctors (CT scans, MRIs, and other diagnostic imaging procedures), and total costs. I find no evidence that doctors change costs or procedure use in response to the program, relative to their behavior at non-participating hospitals. The bonuses create two conflicting forces which may explain why the program did not decrease costs, conditional on patient health. First, there is the intended effect: less care is provided if a patient is admitted under the bonus program than if they were admitted in a hospital with no bonuses. On the other hand, the bonus program causes doctors to admit some patients who otherwise would not have been admitted, and admitted patients receive more care.

While sorting patients between hospitals does not necessarily change treatment costs, it can cause “naive” evaluations of the policy to be biased. In an evaluation of the first wave of the program, the Agency for Healthcare Research and Policy published an article reporting that the bonuses reduced costs per admission by eight percent (AHRQ, Agency for Healthcare Research and Quality (2014)). However, the evaluation only compared the costs of admitted patients at participating hospitals, before and after the program was implemented. My results suggest that a simple pre- versus post- comparison of admitted patients is misleading; it does not take into account changes in admission, nor does it not capture the fact that the composition of admitted patients at these hospitals changed in response to the program.

Related Literature My paper contributes to three main strands of literature. First, it is related to the literature on how doctors respond to financial incentives. There is a large body of work studying how reimbursement levels influence procedure choice, mostly focusing on the decision to perform one particular procedure (see Chapter 3, as well as Clemens and Gottlieb (2014); Coey (2013); Dranove and Wehner (1994); Gruber and Owings (1996); Gruber et al. (1999); Grant (2009); Hadley et al. (2001, 2009); Keeler and Fok (1996); Yip (1998)).³ These papers generally find that doc-

³Most of these papers focus on C-sections, though other procedures such as coronary artery bypass grafting and breast conserving surgery have also been studied.

tors supply more services when payment increases, as well as when the payment of a competing procedure decreases. An implication of this research is that reforms which lower the profit for performing “unnecessary” procedures could be very effective at lowering costs.

Current cost-reduction proposals, however, generally involve changing the entire payment system, which could change doctor behavior on margins other than just procedure choice. To this end, a much smaller branch of the literature has studied how doctors respond to different types of payment systems—for example, fee-for-service versus capitated payments Ho and Pakes (2014); Dickstein (2014).⁴ Unfortunately, studying the effect of payment structure on doctor decision-making is hampered both by data availability, and the fact that doctors practicing under different payment schemes may differ on unobservable characteristics. Therefore, how much and on what margins doctors will respond to payment reform policies remains an open question.

Second, the finding that doctors are able to send healthier patients to participating hospitals is similar to the ability of Managed Care plans to select healthier patients into their plans Duggan (2004); Duggan and Hayford (2013); Leibowitz et al. (1992); Brown et al. (2011). There is much less work, however, on the ability of doctors to identify patients with low expected costs. Doctors selecting patients according to their underlying health has been studied in the context of “report card” policies—public disclosures of the patient health outcomes of individual doctors. The evidence on report cards, however, is mixed; Dranove et al. (2003) find that the introduction of report cards cause cardiac surgeons to select healthier patients, while Kolstad (2013) finds little evidence of selection. Especially with the recent popularity of cost reduction strategies that target doctor pay, it is important to know whether doctors are able to selectively identify low-cost patients to treat.

Third, problems and limitations of pilot programs have been widely studied in economics, particularly in development, education, and environmental economics Duflo (2004); Cullen et al. (2013); Allcott (2015). However, these lessons have generally not been applied to U.S. health

⁴A closely related literature looks at the reaction of *hospitals* to the introduction of prospective payment Cutler (1990, 1995); Ellis and McGuire (1996); Dafny (2005). These papers find that hospitals respond by changing treatment intensity and coding practices in response to DRG specific price changes.

care reform. The Centers for Medicare and Medicaid Services (CMS) has spent millions of dollars on pilot programs (or “demonstrations”) since the 1960s, without considering whether partial equilibrium effects generated by such programs would hold in general equilibrium. Furthermore, the results of these pilot programs help direct the annual spending of Medicare, a 600 billion dollar program. In this paper, I point out that even when there is evidence that such programs are effective, it may be due to gaming rather than true improvements in efficiency.

Roadmap The rest of the paper is organized as follows. Section 1.2 describes the bonus program, and the specific incentives it created for doctors. To formalize the economics behind my empirical results, in Section 1.3 I develop a model of doctor decision-making. I show that the bonuses in the Gainsharing Demonstration clearly incentivize doctors to change who is admitted, and to sort patients between hospitals. The effects of the bonuses on resource use is ambiguous, and so remains a purely empirical question. In the remainder of the paper, I measure the impact of the bonuses empirically. Section 1.4 describes my data and identification strategy, and results are presented in Section 1.5. Section 1.6 concludes.

1.2 Institutional Background

Before describing the details of the pilot program, I first lay out the institutional arrangements of doctors treating patients within a hospital setting. I describe how doctors and hospitals are paid, focusing on Medicare payment rules. Finally, I discuss the New Jersey Gainsharing Demonstration, the details of the bonus calculation formula, and its implementation.

1.2.1 How Doctors Treat Patients within Hospitals

When treating a patient in a hospital, doctors must decide both whether a patient should be admitted, and where to send the patient. Doctors can either decide to admit a patient and treat them, or they can treat the patient in the hospital and then discharge them. The technical definition of

admission is simply that a doctor has written an order of admission. In practice, admitted patients generally stay at least over night and occupy a bed. Within all patient types in my data, there are both patients treated with and without being admitted. When considering whether to admit a patient, a doctor must weigh the benefits against the costs; admitted patients are intensely monitored, and receive more care. On the other hand, admission is costly for the patient, both in terms of time and money. In addition, admitted patients spend more time in the hospital, and thus face a higher risk of contracting a hospital acquired infections, which are often resistant to treatment.

When considering where to treat a patient, doctors are limited to choosing between hospitals where they have pre-arranged relationships. The exact employment relationship between doctors and the hospitals they work within is complicated, and varies from place to place. For the most part, however, doctors treating patients in hospitals are independent contractors, rather than hospital employees. These doctors have arrangements with hospitals which allow them to see patients there—so-called admitting or surgical privileges.⁵ Doctors often have such privileges at more than one hospital (in my data, the average doctor is seen to treat patients at two different hospitals). That doctors tend to work at more than one hospital is key to my identification strategy, as I will compare the behavior of the same doctor that works both in a hospital that offers the bonuses and one that does not.

1.2.2 How Doctors and Hospitals are Paid

Status Quo

For the most part, doctors in the US are paid under the fee-for-service system, and traditional Medicare is no exception. Many argue that this fee-for-service system incentivizes additional care on the margin, and is thought to cause doctors to provide treatments with low or zero marginal benefits. Conversely, hospitals in the US are not paid according to each individual service performed. Instead, hospitals are paid either a fixed amount per visit according to a broad diagnosis category,

⁵Even emergency room doctors are usually not employed by the hospital, but are provided by separate business.

or a per diem for each day spent in the hospital Reinhardt (2006).⁶ Medicare, which makes up approximately a third of the average hospital's net revenue, pays hospitals a fixed sum based on the patient's diagnosis (called diagnosis related groups, or "DRGs")—no matter how expensive the patient is to treat. Thus for Medicare patients, hospitals are incentivized to use fewer services.

The financial incentives of doctors and hospitals over how much care to provide are fundamentally at odds, pushing doctors to do more and hospitals to do less. While hospitals can theoretically constrain doctors' resource use through the threat of revoking their privileges, in reality this is difficult. Doctors are afforded a lengthy due process to protect them from competitive forces that could override quality or patient safety. Furthermore, hospitals benefit from having doctors with privileges on staff, as these same privileges are what bring people into the hospital in the first place. Hospitals would like to use pay incentives to align the incentives of doctors with their own, but it is difficult in the current legal environment. Federal law constrains the ability of hospitals and doctors to participate in cost reduction programs, with the rationale that hospitals will pressure doctors into giving too little care, which would be bad for patient welfare.⁷ Medicare demonstration projects, however, are typically granted waivers to these statutes.

Under the Cost-Reduction Bonus Program

The New Jersey Gainsharing Demonstration was created by the New Jersey Hospital Association to reduce costs by aligning the incentives of doctors with those of hospitals. Under the program, doctors are still paid separately for each service provided, but are now also paid bonuses for lowering costs per visit. These bonuses are paid by the hospital to the doctor, and are supposed to reduce costs by lowering the use of unnecessary procedures. Doctors treating admitted patients at

⁶Medicaid pays hospitals either a flat amount per visit based on diagnosis, or with per diem payments (a lump sum for each day spent in the hospital). Private insurers pay hospitals based on either DRGs, per diems, or discounts negotiated off list charges. Payments from Medicare and private insurers each make up approximately third of hospital revenue Reinhardt (2006).

⁷The civil money penalty (CMP) set forth in section 1128A(b)(1) of the Social Security Act prohibits any hospital or critical access hospital from knowingly making a payment directly or indirectly to a doctor as an inducement to reduce or limit services to Medicare or Medicaid beneficiaries under the doctor's care. In addition, gainsharing arrangements may also implicate the anti-kickback statute (section 1128B(b) of the Social Security Act) and the doctor self-referral prohibitions of the Act (section 1876 of the Social Security Act) Office of Inspector General (1999).

participating hospitals are eligible to receive one bonus per visit, where the maximum bonus they can receive depends on the patient's diagnosis and severity of illness.

When a hospital joins the Demonstration, doctors working in the hospital have the option to sign up for the program. While I do not have on which or how many doctors signed up, anecdotal evidence suggests take-up was high. There is no reason for a doctor to abstain, as there is no change in the process or form of payment, no additional paperwork, and no risk; doctors are only rewarded for improvement, and not punished for stagnation or increasing costs. While many providers are involved in patient care, only the responsible doctor is eligible to receive a cost reduction bonus under the Gainsharing Demonstration. For medical cases, this is the attending doctor, and for surgical cases, it is the surgeon.

The bonus a doctor receives through the Gainsharing Demonstration for treating an eligible (admitted and covered by Medicare) patient is calculated in three steps. First, patients are divided into types based on their diagnosis and how sick they are (for example, one type would be “hip joint replacement, severity of illness level two”), using 3M's All Patient Refined Diagnosis Related Group (APR-DRG) system.⁸ Second, a maximum bonus is assigned to each patient type. All doctors face the same maximum bonus for treating patients of the same type. Third, this maximum bonus is scaled according to whether and how much the doctor reduces costs for that patient type relative to pre-program costs. A hypothetical bonus calculation example is presented in Figure 1.1. In this example, three doctors treat three patients with the same type, but receive different bonuses based on the costs of the treatment they provide.

The maximum bonuses are calculated using cost data from before the program started (the base year was 2007 for the original demonstration and 2011 for the expansion). The maximum bonus for treating a patient type is defined as one tenth of the average deviation from the 25th percentile of the cost distribution for that patient type in the base year. To this end, a third party calculated four maximum bonus amounts for each diagnosis (APR-DRG), depending on the severity of the

⁸As patient types are partially determined by the types and numbers of co-morbidities recorded by the doctor, there is a potential for “up-coding”—doctors changing a patient's diagnosis to increase expected profit. I will discuss this more later in the paper, however, I believe the scope for up-coding is minimal in the Gainsharing Demonstration.

patient's illness (SOI). The four severity of illness categories capture the fact that the same diagnosis (e.g. "peptic ulcer and gastritis") may be more or less serious depending on a patient's age and co-morbidities. I recreate these maximum bonuses using list charges from hospital billing records deflated by Medicare's hospital level cost-to-charge ratio (more details on bonus calculation can be found in the appendix). An example of maximum bonuses for two particular APR-DRGs is given in Table 1.1, and the distribution of maximum bonuses is shown in Figure 1.2.

The maximum bonus, rather than the realized bonus, is the important number when considering the impact of the gainsharing program on doctor behavior. The maximum bonus represents the most a doctor can hope to earn, *ex ante*, for any given patient. A reduction in costs of the same dollar amount for different patient types translates into different realized bonuses, depending on the maximum bonuses. Thus, the size of the maximum bonus reflects how valuable a patient is for participating doctors.

The formulas used to calculate maximum bonuses are based on the idea that high cost variance within a diagnosis is bad, as it suggests that there are high cost patients who could be getting the same treatment as low cost patients. The bonuses are designed to make patients in diagnoses with high cost variance especially profitable. However, the association between cost variance and waste is just a theory. Alternatively, the high cost variance could be due to medical reasons, rather than doctor behavior. If true, diagnoses with high cost variance may be exactly the diagnoses where it is relatively simple to find patients with much lower than average expected costs, making sorting particularly attractive.

1.2.3 Implementation of Gainsharing Demonstration

The Gainsharing Demonstration took place in two waves, which both applied only to admitted Medicare patients.⁹ The initial phase took place in twelve New Jersey hospitals from July 1st,

⁹ During its first incarnation, it was called the New Jersey Gainsharing Demonstration. Later, it was rechristened and expanded as a part of the Bundled Payment for Care Improvement Initiative (BPCI) under the CMS Innovation Center, which was charged by the PPACA to support the development and testing of innovative health care payment and service delivery models. (For ease of exposition, I will call both waves the Gainsharing Demonstration throughout the paper, as the payment incentives were nearly identical.)

2009 to July 1st, 2012. Eight of the original twelve hospitals opted to extend the program through March 31st, 2013. On April 1st, the program was renamed the BPCI Model 1 program, and was expanded to 23 hospitals (including six of the original twelve). My data go from 2006 through the end of 2013, so while I use the variation from the start of the BPCI Model 1 program, the bulk of my variation comes from the first wave of the program, and the 2012 extension.

The hospitals that formed the demonstration and its expansion appear to be similar to other New Jersey hospitals, on average.¹⁰ Figure 1.3 shows that the participating hospitals are scattered around the state, and are thoroughly interspersed with non-participating hospitals (a complete list of participating hospitals can be found in the appendix). As can be seen in Table 1.2, the main difference between participating and non-participating hospitals—especially in the first wave—is that hospitals participating in the program have more Medicare patients on average. That hospitals with more Medicare patients are more likely to participate is to be expected, as hospitals with the most Medicare patients also have the most to gain from a program designed to reduce the costs of treating this population. In addition, hospitals that participated in the first wave were more likely to receive a grade of A on a hospital quality report card. By the second wave, however, these differences disappear. The selection of hospitals into the bonus program is clearly non-random—larger hospitals with more Medicare patients are more likely to participate, and these hospitals may be on different trajectories than non-participating hospitals. However, the identifying variation used in this paper is within doctor, which sidesteps many of the difficulties posed by differential trends at the hospital level.

1.3 Conceptual Framework

To formalize how the bonuses should affect doctor decision-making, I present a stylized model of the incentives and choices faced by doctors working in a hospital setting. The basic setup of the

¹⁰A cap of twelve participating hospitals for the original demonstration was mandated by Medicare, despite considerable interest from additional hospitals. In response, the New Jersey Hospital Association chose the first twelve participants to represent New Jersey hospitals as a whole. As can be seen in Table 1.2, this appears to have been successful.

model is first introduced; I consider a doctor who works in two hospitals, and must decide whether a patient is admitted, where to send the patient, and how much care to provide. I first describe the outcome when neither hospital offers a cost reduction bonus. Next, I introduce the cost reduction bonuses to one of the hospitals in the model. Finally, I compare how the doctor's decisions change as a result of the introduction of the bonuses.

1.3.1 The Set Up

I consider one doctor treating a population of patients with mass one, where all patients are within a single diagnosis-severity of illness cell. I assume for now that the cell is exogenously defined, though I will later examine the validity of this assumption empirically. For each patient, the doctor must make three decisions: whether a patient is admitted, $A \in \{0, 1\}$, which hospital they attend, $H \in \{0, 1\}$, and how much care is provided, $q \in \mathbb{R}^+$. When neither hospital offers a bonus, the two hospitals are identical. Patients vary only by their sickness level β , which is uniformly distributed from zero to $\bar{\beta}$.

Doctors are utility maximizers, and choose H , A , and q to maximize a weighted average of their profit from treating the patient and the patient's utility from treatment, where weight placed on profit is given by λ . Doctors are paid a reimbursement rate, a , for each unit of care, q , provided to the patient. The reimbursement rate, a , does not depend on the hospital choice or whether the patient is admitted. Thus, the doctor's profit from treating a patient is aq .

Doctors do not only maximize profit—they also care about the patient's utility from treatment. A doctor's concern for their patient's welfare can be understood as altruism on behalf of their patients, or as the doctor acting to preserve their reputation. The utility a patient derives from medical care is:

$$\begin{cases} \beta q - \frac{b}{2}q^2 & \text{if } A = 0 \\ \beta q - \frac{b}{2}q^2 + \gamma q - C & \text{if } A = 1 \end{cases} \quad (1.1)$$

where C is a fixed cost of admission. The patient's utility from medical treatment is concave in q , with sicker patients (those with a higher β) benefiting more from medical care. The key

assumption is that patients have a bliss point in q . Care provided past this preferred q need not necessarily become physically harmful, but can be interpreted as patients facing co-insurance and the opportunity cost of their time.

A patient's utility from treatment depends not just on the amount of care provided, but also on whether or not they are admitted. If a patient is admitted to the hospital, there are two opposing effects. On one hand, being admitted makes treatment more beneficial (represented in the model by γ). There are many benefits to being admitted; admitted patients receive more care, and are intensely monitored. On the other hand, the care received by admitted patients is very expensive, and requires a much longer stay in the hospital. The additional care is costly in monetary terms, in terms of a patients' time, and because it translates into a greater probability of contracting a hospital acquired infection. Thus, patients also face a fixed cost of admission, C ; patients dislike being admitted to the hospital, all else equal. When making the decision to admit a patient, a doctor trades off the costs and benefits for their patient, as well as the difference in their compensation.

When doctors are indifferent between hospitals, I assume they randomly assign patients such that they have an equal probability of going to each hospital.^{11,12}

1.3.2 No Bonuses

The two hospitals are identical in the case with no bonuses, and thus the hospital choice drops out—doctors behave the same in each hospital. Doctors are utility maximizers, and choose q and A to maximize a weighted average of their profit from treating the patient and the patient's utility from treatment:

¹¹The randomization can be interpreted as patients having a slight preference for the closest hospital, and patients being evenly distributed across space.

¹²Doctors could assign patients such that any proportion goes to each hospital; I use 50-50 to keep examples simple.

$$\begin{aligned}
\max_{q,A} U(q, A; \beta) &= \underbrace{\lambda [aq]}_{\text{profit}} + (1 - \lambda) \underbrace{\left[\beta q + (\gamma q - C) * \mathbb{1}\{A = 1\} - \frac{b}{2} q^2 \right]}_{\text{patient's utility from treatment}} \\
&= \max \left\{ \underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[(\beta + \gamma) q^*(\beta) - C - \frac{b}{2} q^*(\beta)^2 \right]}_{V_1(\beta) = U(q^*(\beta); \beta, A=1)}, \right. \\
&\quad \left. \underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[\beta q^*(\beta) - \frac{b}{2} q^*(\beta)^2 \right]}_{V_0(\beta) = U(q^*(\beta); \beta, A=0)} \right\}
\end{aligned} \tag{1.2}$$

The intuition is fairly straightforward. Doctors would like to provide as much care q as possible to maximize their profits, but are constrained by patient preferences. Relatively healthy patients dislike admission, while for sicker patients, admission is beneficial. Since doctors take into account patient's preferences, there is a sickness threshold β^A which defines the optimal admission rule.

Proposition 1: Under some parameter conditions, there exists a β^A such that all patients with $\beta < \beta^A$ are not admitted, and all patients with $\beta \geq \beta^A$ are admitted.

The optimal decision rule for admission is depicted in Figure 1.4, which plots the value function of a doctor under two scenarios: all patients being admitted ($V_1(\beta)$), and no patients being admitted ($V_0(\beta)$). Doctors always admit patients when the $V_1(\beta) \geq V_0(\beta)$, and never admit patients when $V_0(\beta) > V_1(\beta)$. β^A is defined as the sickness level where $V_0(\beta) = V_1(\beta)$. Thus, the value function $V(\beta)$ is the upper envelope of $V_0(\beta)$ and $V_1(\beta)$, where the sickest patients are admitted and the healthiest patients are not admitted. As doctors randomize when they are indifferent between hospitals, $\frac{\bar{\beta} - \beta^A}{2}$ patients are admitted at each hospital. A formal proof is presented in the Mathematical Appendix.

1.3.3 With Bonuses

Now, I consider what happens when cost reduction bonuses are introduced at hospital 1. Adding the bonuses only changes the framework described above in one way—doctors' profits change at the bonus hospital:

$$\begin{cases} aq + \max\{\alpha_0 - \alpha_1 q, 0\} & \text{if } H = 1 \text{ and } A = 1 \\ aq & \text{else} \end{cases} \quad (1.3)$$

If an admitted patient is treated at the bonus hospital, the doctor is now eligible to receive a cost reduction bonus: $\max\{\alpha_0 - \alpha_1 q, 0\}$. The bonus is decreasing in the amount of care provided, q , but is never negative. The maximum bonus for the diagnosis-severity of illness group is α_0 , and α_1 represents how quickly the bonus decays as q increases. Everything else remains the same, including the number of patients admitted to the bonus hospital, $\beta' = \frac{\bar{\beta} - \beta^A}{2}$.¹³ Doctors are constrained by the number of patients admitted at the participating in the absence of the bonus program, as the program included language restricting doctors from increasing overall admission. Even if the rules had not mentioned admission levels, holding admission fixed is equivalent to introducing capacity constraints—assuming hospital capacity does not change in response to the program. Doctors can, however, change which patients are admitted and where they are treated. Research has shown that patients essentially do whatever doctors tell them Manning et al. (1987). Since all patients affected by the program are covered by Medicare, and all hospitals accept Medicare, it seems reasonable to assume most patients would agree to use whichever hospital is recommended by their doctor.

Doctors now choose $A \in \{0, 1\}$, $H \in \{0, 1\}$, and q to maximize the utility function

¹³The capacity constraint β' is just a number; doctors can admit any patients they want, and are not constrained to pick patients in an interval of β .

$$\begin{aligned}
\max_{q,H,A} U(q, H, A; \beta) &= \underbrace{\lambda [aq + \max \{\alpha_0 - \alpha_1 q, 0\} * \mathbb{1} \{H = 1, A = 1\}]}_{\text{profit}} \\
&\quad + (1 - \lambda) \underbrace{\left[\beta q + (\gamma q - C) * \mathbb{1} \{A = 1\} - \frac{b}{2} q^2 \right]}_{\text{patient's utility from treatment}} \\
&= \max \left\{ \underbrace{\lambda [aq^*(\beta) + \alpha_0 - \alpha_1 q^*(\beta)] + (1 - \lambda) \left[(\beta + \gamma) q^*(\beta) - C - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_2(\beta) = U(q^*(\beta); \beta, H=1, A=1)}, \right. \\
&\quad \underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[(\beta + \gamma) q^*(\beta) - C - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_1(\beta) = U(q^*(\beta); \beta, H=0, A=1)}, \\
&\quad \left. \underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[\beta q^*(\beta) - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_0(\beta) = U(q^*(\beta); \beta, A=0)} \right\}
\end{aligned} \tag{1.4}$$

subject to the capacity constraint that only $\beta' = \frac{\bar{\beta} - \beta^A}{2}$ patients can be admitted at each hospital. The expression is the same as in the case without the bonus, with the addition of $V_2(\beta)$: the value function if doctors receive the cost reduction bonus.

Whether or not there are bonuses, the admitted patients are always those with the largest (positive) difference between the utility a doctor receives from admitting them and not admitting them. Before the bonuses are introduced, this difference is largest for the sickest patient ($\beta = \bar{\beta}$), and is decreasing in β . The introduction of the bonuses at hospital 1, however, eliminates this monotonicity. The cost reduction bonuses increase the doctor's profit from admitting healthy (low β) patients, up until the point where a patient is sick enough that quantity of care chosen is too high to generate a bonus (represented by the dash-dotted line in Figure 1.5). After the introduction of the bonus, the patients whose admission generates the biggest utility gain are at the extremes: the

lowest β patients because of the bonus, and the highest β patients because these patients have the highest utility from treatment.

Proposition 2: Under some parameter restrictions, there exists a $\tilde{\beta}$ such that patients with $\beta \in [0, \tilde{\beta}]$ are admitted at the bonus hospital, patients with $\beta \in [\tilde{\beta}, \tilde{\beta} + \beta^A]$ are not admitted, and the remaining patients with $\beta \in [\tilde{\beta} + \beta^A, \bar{\beta}]$ are admitted at either the bonus or non-bonus hospital.

After the bonuses are introduced, doctors would like to admit all patients (see Figure 1.5; the upper envelope contains segments of $V_2(\beta)$ and $V_1(\beta)$, but not $V_0(\beta)$). Not all patients can be admitted, however, as doctors are limited by the original hospital capacity—only $\beta' = \frac{\bar{\beta} - \beta^A}{2}$ patients can be admitted at each hospital. The introduction of the bonuses has no impact on the treatment of the sickest patients—doctors will continue to admit them. For the healthiest patients, however, the bonus is large enough that doctors will now admit them, despite the fact that these patients dislike admission. Doctors will admit low β patients at the bonus hospital up until $\tilde{\beta}$. They will also admit the sickest $\bar{\beta} - (\tilde{\beta} + \beta^A)$ patients, randomizing over hospital choice such that they admit β' total patients at each hospital. The patients with β s in the middle of the distribution will not be admitted. This optimal decision rule is shown in Figure 1.5. The exact form of $\tilde{\beta}$, as well as the conditions necessary for an interior solution, are detailed in the Mathematical Appendix.

The cost reduction bonuses introduce two distortions. First, the bonuses increase the probability of admission for the healthiest patients and decrease the probability of admission for sicker patients. Many patients with $\beta < \tilde{\beta}$ are not admitted without the bonus (the “pre-period”), and all are admitted in the when the bonus is introduced (the “post-period”). On the other hand, many “medium sick” patients with $\beta \in [\tilde{\beta}, \tilde{\beta} + \beta^A]$ are admitted in the pre-period, and are not admitted in the post-period. Second, the bonuses cause sorting. After their introduction, doctors send the healthiest patients exclusively to the bonus hospital. Previously, the non-bonus hospital would have received some of the healthier patients, whereas now they only get patients with $\beta > \tilde{\beta} + \beta^A$.

The bonuses’ affect on the quantity of care provided to bonus generating patients, however, is not clear. If a patient is admitted both with and without the bonuses, then q clearly decreases.

If a patient is only admitted under the bonus program, on the other hand, then the change in q is ambiguous. Intuitively, there are two conflicting forces. The first is downward pressure on q from the bonus (represented by α_1). The second is upward pressure on q from admission (represented by γ).

Proposition 3: The direction of the change in q conditional on β from the pre- to the post-period for bonus-generating patients ($\beta \in [0, \tilde{\beta}]$) is ambiguous.

Whether the quantity of care provided for the bonus generating patients is higher or lower than the counterfactual of neither hospital offering a bonus is determined by the relative size of γ and α_1 . For more details, see the Mathematical Appendix.

Finally, the model predicts the results of the naive evaluation. After the bonuses are introduced, the average q for admitted patients falls at the participating hospital. The average q falls because the composition of patients at the participating hospital has changed, not because costs have decreased conditional on patient health (β). A simple comparison of average costs with and without the bonuses, however, would find that costs went down at the participating hospital (see Figure 1.6).

1.4 Data and Empirical Strategy

According to the conceptual framework outlined above, the introduction of the bonus program will cause doctors to change their decisions over admission—both in terms of whether and where patients are admitted. The bonuses will also impact the quantity of services provided, though the direction and magnitude are ambiguous. The relative sizes of these three effects, and whether the program ultimately decreases costs, are empirical questions which I address in the remainder of the paper. In this section, I describe the data and strategy used to investigate each of these questions.

1.4.1 Data Sources

The primary data are the New Jersey Uniform Billing Records, which cover all hospital discharges in New Jersey from 2006 to 2013. Each record in the confidential file includes the patient's name and the medical license number of the attending doctor and surgeon (if the case was surgical). From this raw data, I create a panel by matching patient records across visits by sex, date of birth, and first and last names.¹⁴ I also create doctor identifiers using the recorded license numbers of doctors and surgeons. The final de-identified file includes codes for patients and doctors, allowing me to track them over time and across all hospitals in New Jersey. The ability to follow patients and doctors over time and across hospitals is often lacking in medical records, and is an important strength of this paper. The discharge data also include admission and discharge dates, all diagnoses and procedure codes, payer information, patient demographic information, and list charges. To these data, I add annual information on hospitals from the American Hospital Association (AHA) annual survey, and Medicare's cost-to-charge ratio series.

In addition, I supplement the billing records with the bonus amounts that doctors could have received for treating each patient during the program. While I do not have access to the actual maximum bonuses used, I recreate them based on the formula provided by the New Jersey Hospital Association. The first step is to define the patient types, which is done by passing the billing records through 3M's All Patient Refined - Diagnosis Related Group (APR-DRG) software. For each record, the software creates a diagnosis group, a severity of illness (SOI) category, and designates the visit as medical or surgical. I then combine this information with cost data from the billing records and the bonus formula used in the Gainsharing Demonstration to reconstruct the maximum bonuses that a doctor could earn by treating each patient.

My outcome variables are all constructed from the billing records. To measure how the bonuses affected admission, I create a counterfactual measure for whether a patient would have been admitted in the absence of the program. To measure sorting, I construct measures of each patient's latent

¹⁴The Levenshtein edit distance is used to match names, because of problems with typos and misspellings (Stata command `strgroup`).

health in the year running up to their visit. Resource use and costs measures are taken directly from the billing records—I look at the effect of the bonuses on costs, length of stay, and the use of diagnostic procedures. I will describe each of these outcome variables in more detail in Section 5.

1.4.2 Estimation Sample

The main analysis sample consists of visits where an admitted patient was covered by Medicare, as these are the cases which can generate bonuses for doctors. I restrict the sample to patients seen in general medical and surgical hospitals that were open throughout the sample period. This restriction mainly excludes psychiatric and rehabilitation facilities, which were not targeted by the program. Visits to doctors with very few admitted patients were also dropped, as these doctors likely did not have enough patients to qualify for the bonus program. Finally, my main analysis omits visits where the patient is admitted through the emergency room, as emergency room doctors are not making a hospital choice decision. However, I will come back to patients admitted through the emergency room in a separate analysis at the end.

The main sample includes approximately 400,000 each of medical and surgical visits, which were conducted by 3,474 doctors in 73 hospitals. The patients are predominately white, with an average age of 75, and a high disease burden. The doctors worked at 2.2 hospitals on average, with 35% working in both a participating and non-participating hospital (see Table 1.3). The average maximum bonus a doctor could earn for a surgical patient was \$697, and for a medical patient was \$513. While few doctors take home the whole amount, even receiving half would be a windfall (for comparison, in 2012 Medicare paid doctors \$675.99 to repair a knee ligament Smith, ed (2012)).¹⁵

1.4.3 Empirical Strategy

The main challenge for identifying the effect of the cost-reduction bonuses on doctor decision-making is that participating hospitals are different—and likely on different trajectories—from hospitals that did not take up the program. If participating hospitals are trending differently from

¹⁵Medicare facility charge for repair of knee ligament (CPT 27405), 2012.

non-participating hospitals, comparing the change in outcomes before and after the program was introduced at participating hospitals to non-participating hospitals (a hospital-level difference-in-difference estimator) would be inappropriate. I address this concern by looking within doctor. In this case, the identifying variation comes from choices made by the same doctor working at multiple hospitals. The identifying assumption is now that in the absence of the program, a doctor's behavior would have been on the same trend across all hospitals in which she works.

The regressions will take the form of a generalized difference-in-difference specification with doctor fixed effects:

$$Outcome_{idht} = \beta_0 + \beta_1 Policy_{dht} + \beta_2 X_{idht} + \lambda_t + \lambda_h + \lambda_d + \epsilon_{idht} \quad (1.5)$$

where $Policy_{dht}$ is an indicator for whether the visit occurred in a participating hospital when the bonus program was in effect, and the coefficient of interest is β_1 . The patient characteristics included in X_{idht} vary slightly by specification, but in general contain age, race, and sex, and dummies for patient type. Hospital, quarter, and doctor fixed effects are also included in all regressions (λ_t , λ_h , and λ_d). When I look at the effect of the program on which patients are admitted to the hospital, I further want to know whether patients are differentially affected, depending on the size of the maximum bonus attached to their type. In this case, I interact the policy variable with the bonus size:

$$\begin{aligned} Outcome_{idht} = & \beta_0 + \beta_1 Policy_{dht} + \beta_2 HighBonus_{idht} + \\ & \beta_3 Policy * HighBonus_{idht} + \\ & \beta_4 X_{idht} + \lambda_t + \lambda_h + \lambda_d + \epsilon_{idht} \end{aligned} \quad (1.6)$$

where $HighBonus_{idht}$ is defined as a maximum bonus at or above the median amount across patients. In all specifications, standard errors are clustered at the hospital level.

1.5 Results

All results are presented separately for medical and surgical patients, as there are important differences between these groups. For one, surgical patients have higher admission rates (with a few APR-DRGs at 100% admission), so there is less room to manipulate the admission margin in response to bonuses. Resource use is also higher on average for surgical cases, which is important when I look at length of stay and diagnostics. In addition, the consequences for the patient of changing admission and the quantity of services may be different for medical and surgical cases, which could lead to distinct program effects across the two groups.

1.5.1 The Admission Margin

Outcomes

When considering the effect of the bonuses on admission, I would ideally know whether each admitted patient who was treated under the bonus program would have been admitted if the bonus program did not exist. While I can never know this counterfactual, I estimate it from my data. First, I take data from before the program started, which includes both patients that were admitted and those that were not admitted. Using this pre-program data, I regress admission on a large set of observable characteristics. Next, I use the results of this regression to predict whether or not each patient in my main sample (all of whom are actually admitted) would have been admitted in the pre-period. I call this variable the baseline admission probability, as it answers the question: would a patient with these characteristics have been admitted in the pre-period? If the introduction of the bonus program is associated with a decrease in the baseline admission probability of admitted patients, it is consistent with the bonuses inducing doctors to admit some patients who would not have been otherwise.

Results

Table 1.4 shows the effect of the bonuses on the baseline admission probability. On average, the introduction of the bonuses had no effect, which is consistent both with program rules and binding capacity constraints. While the bonuses are associated with a small decrease in the average baseline admission probability of both medical and surgical patients, it is not statistically significant (seen in Columns 1 and 3 of Table 1.4, for medical and surgical patients respectively). This null result, however, could conceal important heterogeneity with respect to the size of the bonus. While doctors were barred from admitting all of their patients in response to the bonuses, they could change which patients were admitted. As shown in Columns 2 and 4 of Table 1.4, it is important to allow the effect of the policy to differ by the size of the maximum bonus.

For patients in high bonus types, the introduction of the bonuses is associated with a significant decrease in the baseline admission probability—implying that some patients in high bonus types would not have been admitted in the absence of the program. For these patients, the effect of the policy is the sum of the coefficients on the policy and its interaction with high bonus. The policy is associated with a statistically significant decrease of 0.048 percentage points for medical patients (Table 1.4, Column 2) and 0.017 for surgical patients (Table 1.4, Column 4), implying that doctors are more likely to admit patients in high bonus types when the policy is in effect.

Conversely, the bonuses are associated with a small increase in the baseline admission probability of low bonus patients, suggesting that some low bonus patients were not admitted, that would have been admitted in the absence of the program. The effect of the policy on the baseline admission of low bonus patients is simply the coefficient on the policy variable in Columns 2 and 4 of Table 1.4, which is statistically significantly positive in both cases. Low bonus medical patients saw an 0.028 percentage point increase (3% of mean) in the baseline admission rates, and low bonus surgical patients saw an increase of 0.014 percentage points (1.5% of mean). The increase in the baseline admission probability for low bonus patients is consistent with doctors making room for more high bonus patients by not admitting some patients with low bonus types.

The effect of the bonuses on admission is much larger for medical patients, compared to sur-

gical patients. This differential responsiveness is likely due to a combination of two effects. First, admission rates are lower on average for medical patients, leaving more room for discretion. Some surgical APR-DRGs have 100% admission rates, and thus admission cannot increase. Second, surgical procedures are more uniform than medical cases, and have stronger protocols and norms about whether admission is necessary. Thus, it is likely that doctors treating medical patients have more discretion over admission decisions.

1.5.2 The Hospital Sorting Margin

Outcomes

Do doctors sort healthier patients into participating hospitals, conditional on patient type? To answer this question, I need measures of the latent health of patients that are known to the doctor (or at least correlated with information known to the doctor), but not to the bonus formula. My strategy will be to exploit the time-series dimension of the data, and use information from past patient visits.

My preferred measure of latent patient health is the Charlson Co-morbidity Index, which is computed based on 17 conditions weighted by the associated risk of death.¹⁶ This index has been widely validated, and has shown to be strongly predictive of hospital resource utilization Charlson et al. (2008). In order to measure a patient's latent health (rather than the acute event that brought them to the hospital), I construct a "leave-out" version of the Charlson Co-morbidity Index. The leave-out index is constructed using data on each patient's previous visits to the hospital, excluding the current visit. Excluding the current visit helps me identify sorting, as diagnoses from the current visit are taken into account in when determining the severity of illness level, but diagnoses from previous visits are not.

¹⁶The Charlson Co-morbidity Index is a weighted sum over the following 17 conditions, where weights are in parentheses: acute myocardial infarction (1), congestive heart failure (1), peripheral vascular disease (1), cerebrovascular disease (1), dementia (1), chronic pulmonary disease (1), rheumatologic disease (connective tissue disease) (1), peptic ulcer disease (1), mild liver disease (1), diabetes without complications (1), diabetes with chronic complications (1), hemiplegia or paraplegia (2), renal disease (2), cancer (2), moderate or severe liver disease (3), metastatic carcinoma (6), AIDS/HIV (6)

While the Charlson Co-morbidity Index is a useful summary measure of latent patient health, it is by nature incomplete. It only captures a handful of conditions, all of which are very serious. A patient with a high disease burden, but where each individual condition is less serious, may not score highly on the index but still be expensive to treat. Thus, I also look separately at whether patients were treated in the hospital for other medical conditions over the past year: asthma, viral infections, and chronic kidney disease, as well as the number of visits for chronic conditions. Finally, I also look at the total costs generated by hospital visits over the past year, as patients with better latent health should be cheaper to treat at all points in time.

Results

Taking the Charlson Co-morbidity Index as my preferred measure of patient health, I find that doctors admit healthier patients to bonus eligible hospitals in response to the program. Figure 1.7 displays the effect of the bonus policy on the average Charlson Co-morbidity Index of patients in event time (for medical patients), where the implementation of the policy is normalized to $t = 1$. After the policy comes into effect, there is a clear drop in the average co-morbidity burden of patients.

The sorting result depicted in Figure 1.7 is presented in regression form in Column 1 of Table 1.5. The bonuses are associated with a decrease in the average Charlson Co-morbidity Index of medical patients of 0.11. Medical patients admitted in hospitals offering bonuses are significantly healthier. The patient sorting channel is quantitatively substantial; using estimates from Quan et al. (2011), a decrease of 0.11 is associated with decreasing in-hospital mortality by 5 - 18%. In addition, doctors have shown that the Charlson Co-morbidity Index is strongly predictive of resource utilization Charlson et al. (2008), suggesting that these healthier patients are indeed cheaper to treat. The Charlson Index of surgical patients does not respond to the program, probably because the diagnoses included in the index are more closely tied to medical conditions than surgical ones (Appendix Table 1.11 lists the top 15 primary diagnoses for medical and surgical patients in my sample).

In addition, the number of chronic conditions recorded, as well as the probability patients were seen for asthma, viral infections, or chronic kidney disease in the past year all decreased for medical patients admitted in participating hospitals (Columns 3 - 5 of Table 1.5). These patients also have accumulated fewer hospital costs over the past year (Column 2 of Table 1.5), though the estimate is not statistically significant. The effects for analogous surgical patients are much noisier, though they are less likely to have been seen for asthma over the past year.

Exactly as suggested by the model, Table 1.5 shows that conditional on type, patients admitted by doctors at participating hospitals are healthier than patients admitted by the same doctors at non-participating hospitals. The mechanism is straightforward—conditional on diagnosis, healthier patients are cheaper to treat, and cheaper patients earn higher bonuses. By sorting patients across hospitals, doctors can earn a bonus without changing treatment conditional on admission and patient health.

One might wonder whether the results in Table 1.5 are driven by changes in the admission margin—patients that are only admitted because of the program are likely healthier on average—rather than pure sorting. Table 1.6 addresses this concern by repeating the analysis on the subsample of patients in diagnoses that are nearly always admitted, as it is unlikely that the composition of this sample would be affected by changing admission thresholds. While the results are less precisely estimated, the magnitudes are similar. Thus, it appears that doctors are responding to the bonuses by both changing admission thresholds and sorting patients across hospitals.

Not only do doctors sort healthier patients into participating hospitals, the hospitals that participated in the first wave are also higher quality—or at least score better on one particular hospital quality report. In general, if the participants in pilot programs are higher quality, another unintended consequence of programs which incentivize doctors and hospitals to cream-skim patients could be a worse match between patient health and treatment quality.

Up-Coding

Both the model and the sorting results assume that the diagnosis and severity of illness margins are unaffected by the Gainsharing Demonstration. One might be worried about this assumption, as during the 1990s many hospitals were accused of up-coding—exaggerating a patient’s diagnosis to extract a higher reimbursement from Medicare. Silverman and Skinner (2004) found, for example, that between 1989 and 1996, the percentage point share of the most generous diagnosis groups (DRGs) for pneumonia and respiratory infections rose precipitously. One reason the diagnosis groups used by Medicare are particularly susceptible to up-coding is that there are often multiple DRGs for each diagnosis, where the most severe version pays a much higher amount. For example, there are separate Medicare diagnosis groups (MS-DRGs) for diabetes with complications (638), diabetes with major complications (637), and diabetes without complications (639), where the more severe codes are reimbursed at higher rates. In the diagnosis groups used for the bonus calculations, however, this feature is lacking. In order to upcode at the diagnosis level doctors would have to change the diagnosis conceptually, which seems unlikely (e.g., changing a diagnosis from “diabetes” (APR-DRG 420) to “malnutrition, failure to thrive, and other nutritional disorders” (APR-DRG 421)).

It is also possible that doctors could respond to the Gainsharing Demonstration by trying to move their patients into higher severity of illness bins.¹⁷ Influencing the severity of illness (SOI) designation should be difficult, as it is imputed by software and not recorded by the doctor. The only way doctors can affect the severity of illness is to change which secondary diagnoses are recorded on a patient’s chart. While the link between any one co-morbidity and the designation generated by the software is not clear, adding additional diagnoses to all patients could lead to higher average SOI designations. If doctors managed to inflate the severity of illness of all patients in response to the program, the average “true sickness level” of the patients in each cell would decrease—the sickest patients in the first severity bin would be shifted into the next bin, and so on

¹⁷Though while a “with complications” designation always leads to a higher payout in the Medicare DRG system, a higher SOI level does not necessarily lead to a higher bonus.

up the chain. Up-coding, therefore, could generate similar patterns in the data as sorting.

There is no association between the bonuses and the average severity of illness, however, suggesting that up-coding is not a concern in this context (see Table 1.7). The regressions reported in Table 1.7 use the same empirical strategy outlined in Equation 1.5, but with diagnosis rather than diagnosis by severity of illness fixed effects. One interpretation of this null result is that hospitals are able to closely monitor the coding practices of their doctors. The proximity of the payor (the hospital) to the recipient (the doctor) in the Gainsharing Demonstration differs substantially from earlier settings where up-coding has been found. Even if doctors are able to influence the severity codes, it may be much harder to upcode patients when working within the walls of the entity making the payment, in comparison to a distant third party such as Medicare.

1.5.3 The Quantity of Services Margin

Outcomes

To examine whether the bonuses changed procedure use or costs, I again use a variety of measures. The first two are summary measures of resource use: how long a patient stays in the hospital, and the total costs incurred. Length of stay is defined as the number of nights spent in the hospital, and is often used to proxy for the intensity of care provided during the visit. The total costs incurred during a visit are estimated using the total list charges reported in the discharge data, deflating them by Medicare's hospital-year level cost-to-charge ratio, and then converting them to real 2010 dollars. Deflating the list charges is an important step, as list charges are closer to bargaining tools than a measure of the costs to the hospital of providing a service. The cost-to-charge ratio files, however, are explicitly designed to translate list charges into an estimate of the resource cost of inpatient care.

In addition to summary measures of resource use, I look specifically at the use of diagnostic imaging to proxy for the use of unnecessary procedures. While it is difficult to pinpoint any specific test as unnecessary, there is widespread agreement that diagnostic imaging is overused Hillman and

Goldsmith (2010); Abaluck et al. (2015).¹⁸ Thus, if the bonuses are associated with a reduction in use of expensive diagnostic imaging procedures such as magnetic resonance imaging (MRI) and computed tomography (also called CT or CAT scans), it would be consistent with the bonuses lowering the use of unnecessary procedures. The bonuses could also cause doctors to substitute expensive tests for cheaper tests; in particular, I look at whether the bonuses increase the use of diagnostic ultrasounds, which are cheap and radiation-free imaging tests.¹⁹

Results

As shown in Table 1.8, the bonus policy has no significant effect on costs or resource use, conditional on admission and patient health—despite the fact that the program was designed to reduce costs. The program is not associated with significant decreases in length of stay, diagnostic imaging tests, or costs. Even taking the point estimates at face value, the magnitudes are small, and the signs are not even consistently negative. I also find no evidence of substitution between high-tech (MRIs and CT scans) and low-tech (diagnostic ultrasounds) imaging, consistent with the disappointing results of the Medicare Imaging Demonstration, which tried to reduce inappropriate use of high-tech imaging through decision support software Timbie et al. (2014).

Furthermore, while the regressions reported in Table 1.8 include controls for all latent health measures examined in the previous section, it is likely that I cannot completely control for differences in underlying health status. Given the fact that healthier patients were sorted into participating hospitals sorting, these patients may have required fewer resources from the start. Thus, the small decreases reported in some measures in Table 1.8 should be considered an upper bound on the true effect.

Given these results, how did the naive evaluation conclude that the program succeeded in decreasing costs? In Table 1.9, I attempt to replicate the naive evaluation of the first wave of the

¹⁸For example, over half of the procedures labeled by doctors as unnecessary in the *Choosing Wisely* campaign (<http://www.choosingwisely.org/>) are directly related to diagnostic imaging Rao and Levin (2012)

¹⁹Unnecessary diagnostic imaging not only contributes to high health care costs—it may also harm patients. False positives can lead to additional treatments with much higher health risks. With CT scans there is also a risk that patients will react to the contrast material, which is rare but serious Lessler et al. (2010). In addition, radiation exposure may increase later cancer risk Smith-Bindman (2010).

program. In Column 1, I only include hospitals that eventually take up the program, with no health controls or doctor fixed effects. Here, the policy appears to decrease costs, and this decrease is statistically significant. In Column 2, however, I show that clustering standard errors at the hospital level renders the decrease in costs insignificant. In Columns 3 - 5, I add comparison hospitals, health controls, and doctor fixed effects, and show that the sign flips from negative to positive. An incomplete analysis can conclude the program lowered costs, but this conclusion does not hold in to a more thorough investigation.

1.5.4 Extensions and Robustness Exercises

Placebo tests

To confirm that my results are not spurious, I conduct two placebo tests. First, I randomly assign New Jersey hospitals to participate in the program, holding constant both the number of participating hospitals and the timing. I repeat my main regressions using randomly assigned participation, and plot the CDFs of the resulting coefficients (based on 100 repetitions) in Appendix Figure 1.11. Second, I hold fixed the true hospital participation, but randomly assign start dates for the program, and again repeat all of my main regressions. The CDFs of the coefficients from this second simulation (again based on 100 repetitions) are presented in Appendix Figure 1.12. The coefficients from the true regressions are represented by a vertical line, and the 90th percentile by a horizontal line. In nearly all cases, the true coefficients are well above the 90th percentile—in a few cases, the true coefficients are larger than any coefficient generated under the simulation. The results of the simulations suggest that it is extremely unlikely that my findings are due to chance.

Other patient groups

Emergency Room Medicare Patients While patients admitted through the emergency room were excluded from the main analysis, they provide both a useful placebo test for sorting, and another group of patients among which costs might have been reduced. Patients admitted through

the emergency room (ER) cannot be sorted in response to the bonuses, as emergency room doctors cannot send a healthier than average ER patient to a different hospital. In Appendix Table 1.13, I show that the sorting results pass this placebo test—the introduction of the bonuses is not associated with a change in underlying patient health for patients admitted through the emergency room.

In addition, the doctor who decides to admit a patient from the emergency room is not necessarily the doctor who would receive the bonus. While it varies from hospital to hospital, an attending in the emergency department generally decides to admit a patient, and then a different doctor is responsible for the patient after admission. The way the Gainsharing Program is designed, a bonus from treating an ER patient would most likely go to the second doctor, and thus the bonuses should not influence admission. This intuition is consistent with the data, as there is no effect of the bonus program on baseline admission for patients admitted through the ER (see Appendix Table 1.14).

Exactly because there is no sorting margin and a limited admission margin, however, the bonuses could have a large effects on costs and procedure use for patients admitted through the emergency room. In Appendix Table 1.15, however, I show that the effects of the bonuses on costs and quantity measures for these patients are similar to those in the main sample. While the point estimates for the effect of the policy on length of stay is larger for emergency room patients than in the main sample, they are not statistically significant.

Patients Not Covered by Medicare Despite the fact that only Medicare patients are included in the bonus program, it is possible that the program could spill over into the treatment of other patients—particularly “near Medicare” patients. “Near Medicare” patients, aged 50 to 64, have many of the same health problems as the Medicare population but are too young to qualify for Medicare coverage.²⁰ The bonus program could spill over into the care of near Medicare by crowding out care for near Medicare patients, as their treatment cannot generate a bonus.

I find no evidence that the program caused doctors to crowd out younger patients (see Appendix

²⁰Medicare patients who are younger than 65 are excluded from this analysis. Patients younger than 65 can be enrolled in Medicare, for example those on disability or with end-stage renal disease.

Tables 1.16 through 1.18). If anything, the results for near Medicare patients point in the same direction as the main results, though are roughly 80% smaller in magnitude. While any treatment spillovers are very small, these results are consistent with it being difficult for doctors to perfectly target Medicare patients—that some patients who are not eligible for the program still end up being treated like Medicare patients.

Patients of Single Hospital Doctors Finally, I consider the subsample of hospital visits to doctors who only practice at one hospital. These “single hospital doctors” are interesting for two reasons. First, they cannot move patients between hospitals based on their latent health, and so this sample provides an additional placebo check for the sorting results. Second, the doctors that work in only one hospital provide insight into what we might expect if the program was scaled up. If all hospitals participated in the program, the sorting margin would be shut down, but doctors could still respond on both admissions and costs. As doctors that work in just one hospital face similar incentives (they can change admissions and costs, but cannot sort patients between hospitals), they are a useful comparison group. However, using single hospital doctors in a hospital-level difference in difference specification, I find no evidence that the program becomes more effective at reducing costs when the sorting margin is shut down (see Appendix Tables 1.19 through 1.21).²¹

Alternative Strategy: Simulated Share Treated

The doctor-level difference-in-difference specification paints a clear picture: doctors respond to the bonuses by changing their behavior in participating hospitals, relative to non-participating hospitals: they manipulate admission and sort patients to maximize their bonuses, but do not reduce costs. Policymakers, however, may additionally want to know what effect the bonuses had on total costs, rather than relative costs. I use an alternative identification strategy to examine the effect of the program on total costs incurred by doctors, which asks whether ex-ante program exposure affects doctor-level resource use.

²¹The single hospital doctor analysis uses a hospital-level difference in difference strategy, which is the same as the main specifications, but without doctor fixed effects.

In order to isolate the effect of the bonuses on total costs net of sorting, I create a doctor-level exposure variable: the simulated share of a doctor's case load that is treated. Specifically, the simulated share is the fraction of a doctor's patients that would be affected by the program if the distribution of patients across hospitals was fixed in the pre-period. The simulated share is zero in the pre-period, and then rises to the fraction of a doctor's caseload treated at participating hospitals before the program started (2006-2008). The simulated share treated captures the fact that some doctors only admit patients to participating hospitals (their simulated share goes from zero to one), others are not exposed at all, and many doctors are in between. By construction, the simulated share reflects only ex ante exposure, and will not be affected by doctors sorting patients in response to the bonus program.

I collapse the data to the doctor-quarter level, and regress the simulated share treated regressed on cost and quantity measures:

$$outcome_{dt} = \beta_0 + \beta_1 share\ treated_{dt} + \lambda_d + \lambda_t + \epsilon_{dt} \quad (1.7)$$

where λ_d and λ_t are doctor and quarter fixed effects. The coefficient of interest is β_1 , which I interpret as the effect of program exposure on total costs, net of sorting.

There is no evidence of any cost-saving response to the Gainsharing Demonstration in response to program exposure, backing up the main results. As shown in Column 1 of Table 1.10, if anything, exposure to the Gainsharing Demonstration leads to an increase in costs (statistically significant at the 10 percent level). This result concords with the main results of the paper, and could be a result of lowering the threshold for admission (the effect of program exposure on the number of admitted patients is positive, but insignificant). There is no affect of the bonuses on the use of CT scans or MRIs, and a small positive affect on the use of diagnostic ultrasounds—also suggestive of a change in the composition of admission towards healthier patients. Overall, there is no evidence from either identification strategy that the Gainsharing Demonstration resulted in lower costs.

1.6 Conclusion

In this paper, I show that a pilot program that paid doctors bonuses for reducing costs was unsuccessful; doctors changed which patients were admitted and sorted healthier patients into participating hospitals, but did not reduce costs. That doctors are able to identify low-cost patients even within risk-adjusted groups has ramifications beyond the success or failure of this particular program—it is crucial information for policy design. Furthermore, for pilot programs to be informative, they must contain the same incentives as if the program was expanded nationwide. If doctors are able to sort patients in and out of pilot programs in response to changing incentives, this assumption fails.

The results of this paper are applicable to many active programs experimenting with doctor incentives. For example, the Affordable Care Act has heavily promoted Accountable Care Organizations (ACO) as an improvement on the fee-for-service system. ACOs are networks of doctors and hospitals who share responsibility for providing care to patients in an attempt to limit duplication of services and other unnecessary spending. Like Gainsharing, ACOs create incentives for savings by offering bonuses when providers reduce costs. Also like Gainsharing, ACOs only cover some patients, and doctors can influence which patients are covered by the program. In future work, I will investigate whether doctors participating in ACOs exhibit similar sorting behavior to those in the New Jersey Gainsharing Demonstration.

The Department of Health and Human Services (HHS) recently committed to moving 50 percent of Medicare payments (and 90 percent of fee-for-service Medicare payments) into “value-based” payment models by 2018. Many alternative payment models will be used to meet the 2018 goal—such as “Accountable Care Organizations (ACOs), advanced primary care medical home models, new models of bundling payments for episodes of care, and integrated care demonstrations for beneficiaries that are Medicare-Medicaid enrollees” [HHS, Department of Health and Human Services (2015)]. The early evidence on whether these alternative payment models seem to work comes largely from pilot programs and other partial equilibrium settings. However, my paper calls into question the ability of pilot programs to provide this type of evidence, due to the

sorting and selection behavior of providers. Furthermore, incomplete expansion of these payment models continues to leave room for providers to act strategically in ways that undermine the reform's success.

1.7 Figures

Figure 1.1: Hypothetical Bonus Calculation

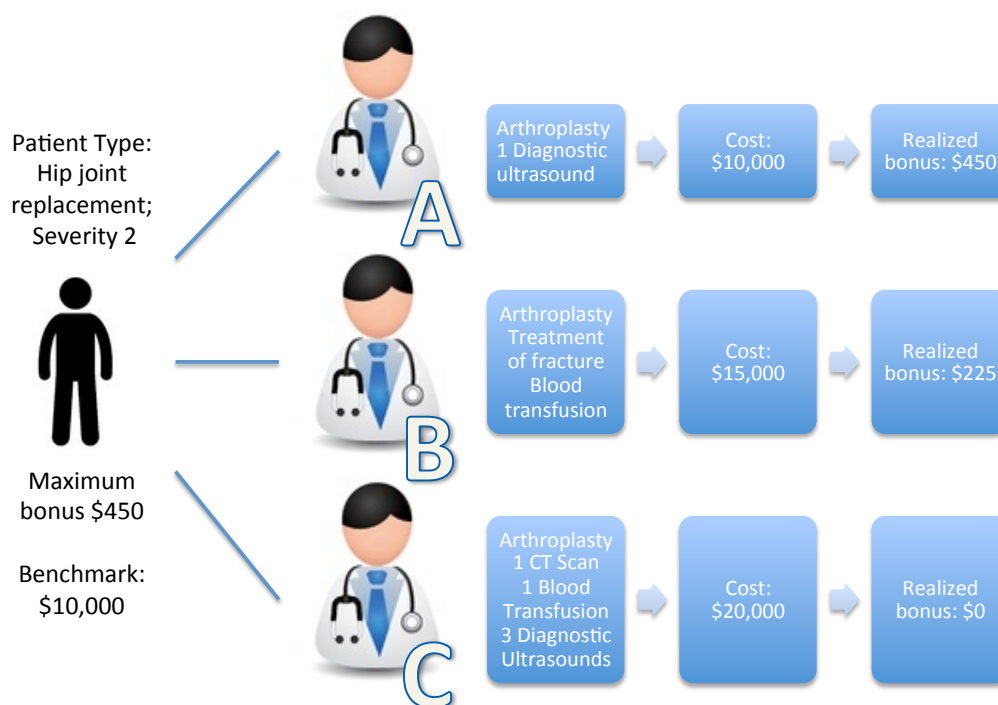
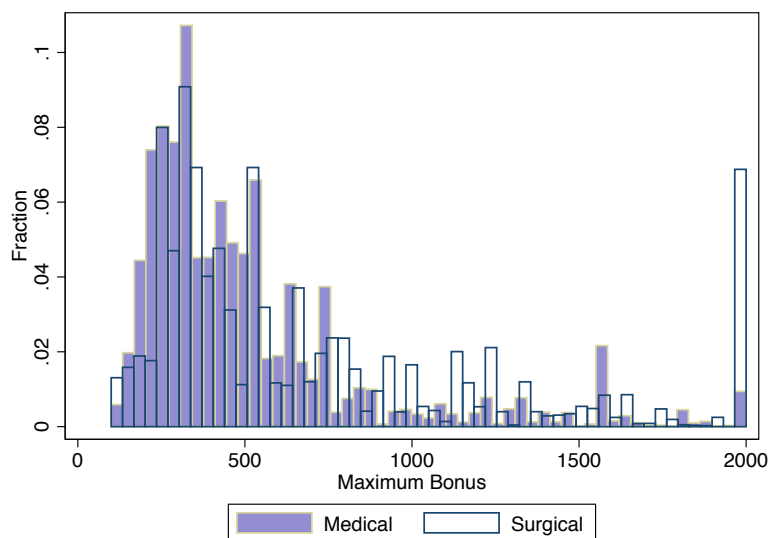
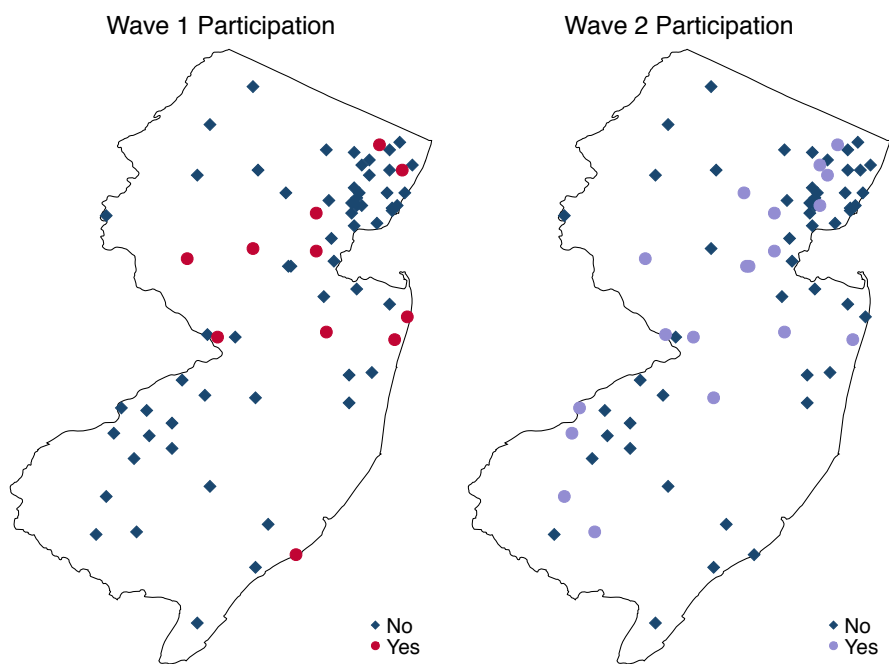


Figure 1.2: Distribution of Maximum Bonuses



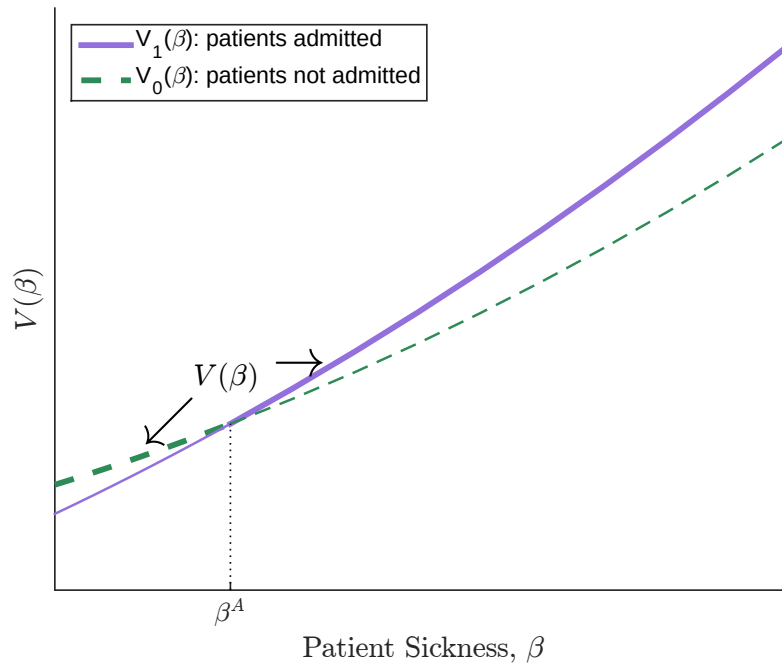
Notes: Each observation is a Medicare beneficiary's inpatient visit to a general medical/surgical hospital in New Jersey from 2006-2013, excluding visits that went through the emergency room.

Figure 1.3: Hospital Locations



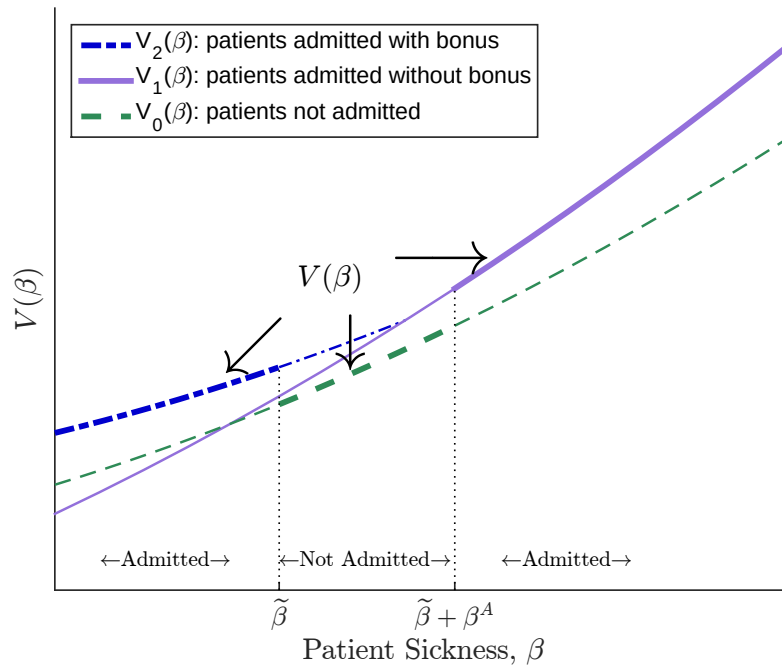
Notes: Diamonds are hospitals that never participated, circles are hospitals that took up the bonuses in the first wave (on the left) and the second wave (on the right).

Figure 1.4: Doctor's Utility as a Function of β : without Bonuses



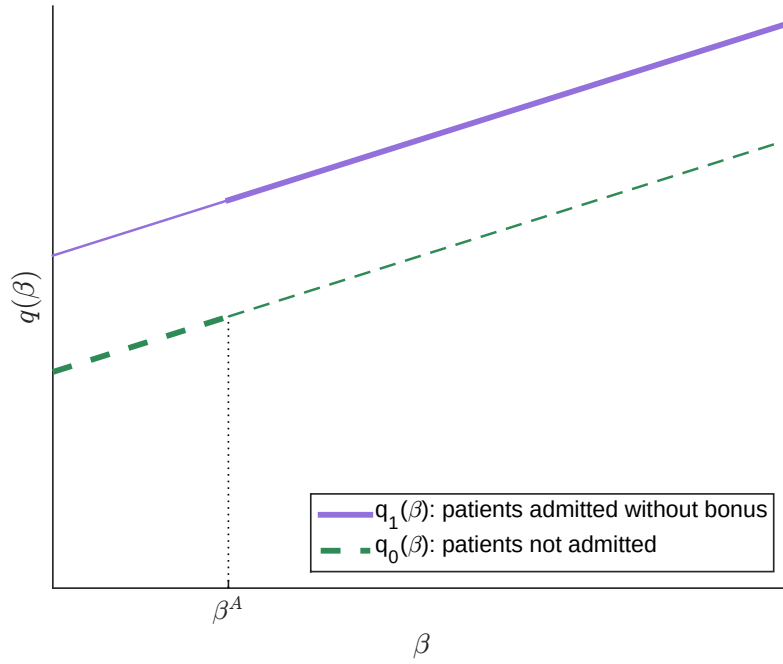
Notes: The bold line sections show the optimal decision rule as a function of β .

Figure 1.5: Doctor's Utility as a Function of β : with Bonuses

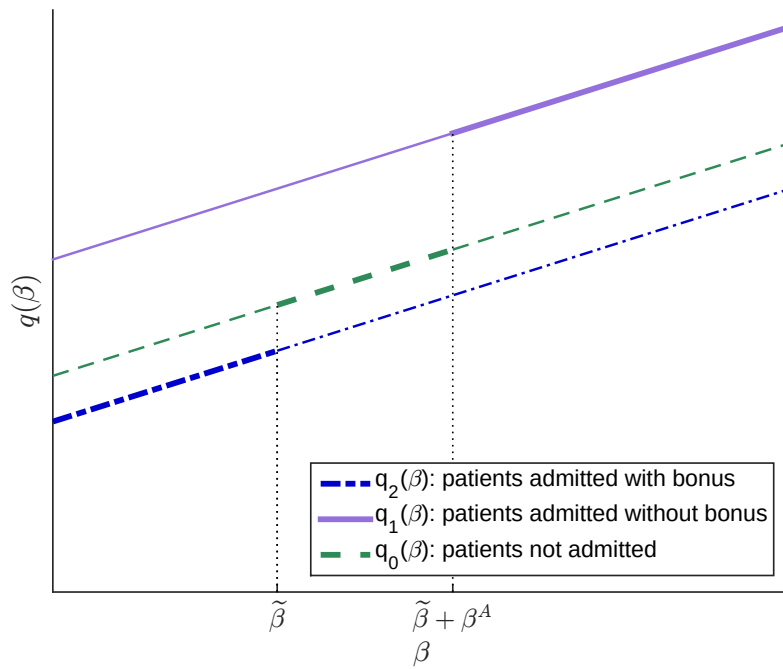


Notes: The bold line sections show the optimal decision rule as a function of β .

Figure 1.6: Optimal Quantity of Care as a Function of β
Without Bonus

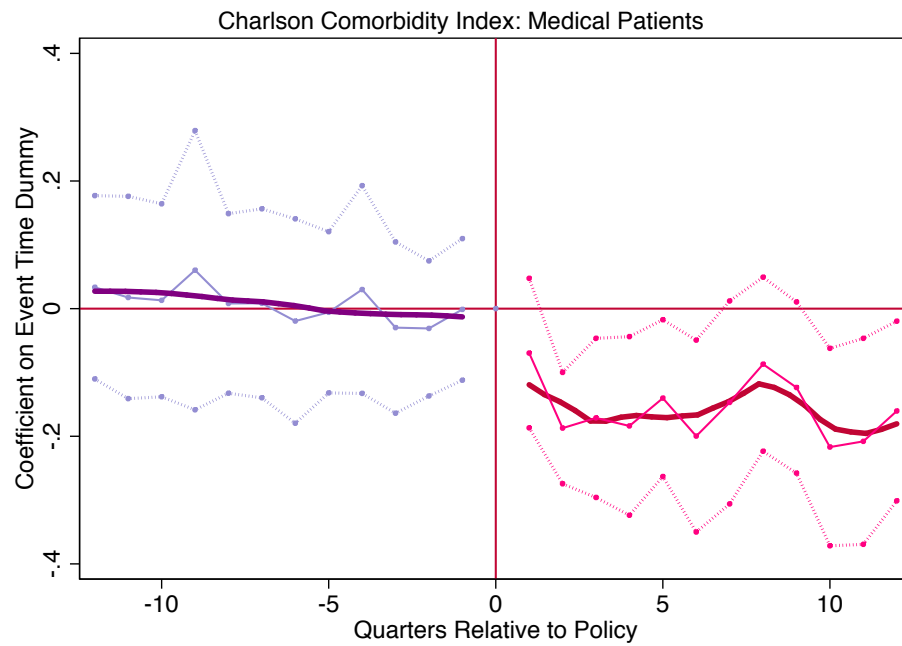


With Bonus



Notes: The bold line sections show the optimal quantity of care provided as a function of β , both with and without bonuses.

Figure 1.7: Healthier Patients Sent to Participating Hospitals



Notes: Each observation is a Medicare beneficiary's inpatient visit to a general medical/surgical hospital in New Jersey from 2006-2013, excluding visits that went through the emergency room.

1.8 Tables

Table 1.1: All Patient Refined-Diagnosis Related Group (APR-DRG) Examples

All Patient Refined-Diagnosis Related Group (APR-DRG)	Severity of Illness (SOI)	Maximum Bonus	Number of Patients
Peptic Ulcer and Gastritis	1	\$189	632
	2	\$280	1,871
	3	\$510	1,552
	4	\$1,4034	317
Hip Joint Replacement	1	\$308	15,711
	2	\$433	12,341
	3	\$911	1,439
	4	\$1,669	557

Notes: APR-DRG and SOI from 3M's grouping software; maximum incentive calculated according to gainsharing formula. Number of patients are for admitted Medicare patients in main sample, which excludes those who went through the emergency room.

Table 1.2: Hospital Characteristics

	Wave 1			Wave 2		
Participation	No	Yes	Diff	No	Yes	Diff
# of Hospitals (Gen. Medical/Surgical)	53	12		42	23	
Nongovernment Not -for-Profit	0.83	0.83	0.00	0.77	0.9	-0.13
Bed Size Code	5.40	6.00	-0.60	5.40	5.60	-0.20
ER Visits	51,194	54,469	-3,275	48,701	55,411	-6,710
Hospitals in a Network	0.54	0.58	-0.04	0.52	0.59	-0.07
CBSA Type: Metro (Pop. of 50,000+)	0.17	0.17	0.00	0.17	0.17	0.00
Medicare Discharges	6,393	8,449	-2,056*	6,236	7,400	-1,164
Medicare Days	39,007	50,011	-11,004	37,999	44,585	-6,586
Medicaid Discharges	2,445	1,950	495	2,365	2,341	24
Medicaid Days	12,085	8,575	3,510	11,558	11,295	263
“Grade A”	0.40	0.67	-0.27*	0.45	0.44	0.01

Notes: Data from the 2008 American Hospital Association Annual Survey—the year before the program was implemented. Medicaid/Medicare days are the total number of hospital days used by people with these insurers. Grade A refers to hospitals that scored an “A” on their hospital report card, as reported by the Leapfrog Group (<http://www.hospitalsafetyscore.org/>). *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.3: Main Sample Characteristics

Patients		Doctors		Outcomes	
N	815,014	N	3,474	Charlson index	2.26
N (medical)	389,604	Avg. # of patients	215	Total chronic	3.52
N (surgical)	425,410	Avg. # of hospitals	2.2	Viral Infections	0.02
Avg. age	74	Med. # of hospitals	2.0	Kidney disease	0.10
% white	81	Ever in policy hosp	74%	Asthma	0.03
% black	11	Ever in other hosp	39%	Length of stay	6.74
% woman	54	Ever in both types	35%	CT scans	0.03
% in policy hosp	19	Avg. max bonus (med.)	\$513	MRIs	0.01
APR-DRGs (med.)	163	Avg. max bonus (surg.)	\$697	Diag. imaging	0.15
APR-DRGs (surg.)	117				

Notes: Statistics for main sample of admitted Medicare patients in New Jersey general medical and surgical hospitals, from 2006-2013, which excludes visits that went through the emergency room. APR-DRG stands for All Patient Refined - Diagnosis Related Group. % under policy refers to the percent of patients treated at participating hospitals while the program was in effect.

Table 1.4: Bonuses Change Which Patients are Admitted

	Medical Patients		Surgical Patients	
	(1)	(2)	(3)	(4)
	Baseline	Baseline	Baseline	Baseline
	Admission	Admission	Admission	Admission
policy	-0.004 (0.006)	0.028*** (0.005)	-0.008 (0.006)	0.014*** (0.004)
high bonus		0.043*** (0.012)		-0.027*** (0.005)
policy * high bonus		-0.076*** (0.006)		-0.031*** (0.005)
Doctor FEs	✓	✓	✓	✓
Mean	0.90	0.90	0.95	0.95
Clusters	73	73	73	73
<i>N</i>	385,845	385,845	405,400	405,400

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. Standard errors clustered at the hospital level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.5: Doctors Sort Healthier Patients into Participating Hospitals

Panel A: Medical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Charlson	Tot. Costs	Tot. Chronic	Asthma	Viral Inf.	Chron. Kidney
policy	-0.111** (0.042)	-701.617 (524.629)	-0.076* (0.047)	-0.004** (0.002)	-0.002** (0.001)	-0.008*** (0.003)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	2.785	17,916.9	4.167	0.0401	0.0188	0.115
Clusters	73	73	73	73	73	73
<i>N</i>	389,604	389,604	389,604	389,604	389,604	389,604

Panel B: Surgical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Charlson	Tot. Costs	Tot. Chronic	Asthma	Viral Inf.	Chron. Kidney
policy	0.013 (0.033)	128.350 (219.007)	0.024 (0.058)	-0.002** (0.001)	-0.001 (0.001)	0.001 (0.003)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	1.775	9,048.8	2.917	0.0266	0.00948	0.0848
Clusters	73	73	73	73	73	73
<i>N</i>	425,410	425,410	425,410	425,410	425,410	425,410

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. CCI stands for Charlson Co-morbidity Index, which is calculated based on information in previous visits. Tot. Chronic refers to the number of body systems affected by chronic conditions. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.6: Doctors Sort Healthier Patients into Participating Hospitals: High Admission Diagnoses

Panel A: Medical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Charlson	Tot. Costs	Tot. Chronic	Asthma	Chron. Kidney	Viral Inf.
policy	-0.127** (0.057)	-528.816 (706.427)	-0.050 (0.064)	-0.003 (0.003)	-0.001 (0.002)	-0.008** (0.004)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	3.103	21,806.9	4.530	0.0461	0.0206	0.123
Clusters	73	73	73	73	73	73
N	162,686	162,686	162,686	162,686	162,686	162,686

Panel B: Surgical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Charlson	Tot. Costs	Tot. Chronic	Asthma	Chron. Kidney	Viral Inf.
policy	0.005 (0.027)	244.348 (164.893)	0.016 (0.064)	-0.003* (0.002)	-0.002* (0.001)	-0.002 (0.002)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	1.242	5,108.9	2.269	0.0240	0.00758	0.0494
Clusters	71	71	71	71	71	71
N	127,163	127,163	127,163	127,163	127,163	127,163

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. CCI stands for Charlson Co-morbidity Index, which is calculated based on information in previous visits. Tot. Chronic refers to the number of body systems affected by chronic conditions. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.7: Effect of Program on Population-Level Severity of Illness

	Medical		Surgical	
	(1)	(2)	(3)	(4)
	SOI	SOI	SOI	SOI
	All	High Adm.	All	High Adm.
policy	-0.026	-0.024	0.011	0.015
	(0.021)	(0.028)	(0.014)	(0.021)
Doctor FEs	✓	✓	✓	✓
Mean	2.363	2.560	1.996	1.945
Clusters	73	73	73	71
<i>N</i>	389,604	162,686	425,410	127,163

Notes: Quarter, doctor, hospital, and diagnosis fixed effects also included, as well as dummies for age categories, sex, and race. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.8: Bonuses Do Not Reduce Costs or Change Procedure Use

Panel A: Medical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Diag. Ultra	Any Imaging	Total Costs
policy	-0.093 (0.196)	-0.000 (0.003)	-0.001 (0.002)	-0.001 (0.003)	-0.002 (0.006)	391 (516)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	6.979	0.032	0.016	0.037	0.112	11,459
Clusters	73	73	73	73	73	58
<i>N</i>	389,604	389,604	389,604	389,604	389,604	355,306

Panel B: Surgical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Diag. Ultra	Any Imaging	Total Costs
policy	-0.042 (0.100)	-0.003 (0.003)	-0.000 (0.001)	0.001 (0.006)	0.008 (0.010)	889 (560)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	6.516	0.025	0.008	0.061	0.186	18,557
Clusters	73	73	73	73	73	58
<i>N</i>	425,410	425,410	425,410	425,410	425,410	371,262

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race, and the variables measuring underlying health from Table 1.5. Standard errors clustered at the hospital level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.9: Replicating the Naive Evaluation: Medical Patients

	(1)	(2)	(3)	(4)	(5)
	Total Costs	Total Costs	Total Costs	Total Costs	Total Costs
policy	-469** (219)	-469 (869)	-364 (849)	414 (636)	606 (612)
Health Controls			✓	✓	✓
Comparison Hospitals				✓	✓
Doctor FEs					✓
Mean	10,910	10,910	10,910	11,281	11,281
Clusters	-	11	11	58	58
<i>N</i>	85,374	85,374	85,374	293,052	293,052

Notes: Quarter, hospital, and diagnosis by severity of illness fixed effects also included. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.10: Simulated Share Treated on Costs and Procedure Use

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Total Costs	Patients Admitted	CT Scans	MRIs	Diagnostic Ultrasounds	Any Diag. Imaging	Total LOS
Simulated share	8,432.72* (4561.81)	0.25 (0.38)	0.08 (0.07)	-0.01 (0.03)	0.08** (0.04)	0.22** (0.10)	1.48 (2.39)
Doctor FEs	✓	✓	✓	✓	✓	✓	✓
Mean	210,940.9	19.99	1.078	0.437	0.835	2.862	122.8
Clusters	3,394	3,394	3,394	3,394	3,394	3,394	3,394
<i>N</i>	84,626	84,626	84,626	84,626	84,626	84,626	84,626

Notes: Doctor and quarter fixed effects included. The outcome variables are summed at the doctor quarter level. Total LOS is the sum of length of stay across a doctor's patients in each quarter. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

1.9 Appendix

Bonus Calculation Details:

Maximum Bonus:

The maximum bonus is calculated using cost data from 2007, before the program started. Within each diagnosis and severity of illness level pair, the maximum bonus is ten percent of the average deviation of costs from the 25th percentile of costs:

$$0.1 * \left(\frac{1}{n} \sum_{i=1}^n (c_i - c_{25^{th}pctile}) \right) \quad (1.8)$$

where c_i is the cost of care for a patient in 2007 (before the program), and $c_{25^{th}pctile}$ is the 25th percentile of the cost distribution for the particular diagnosis/severity pair in 2007. The maximum bonus then is constrained to be between \$100 and \$2000.

I calculate these maximum bonuses using all inpatients over 55 at general medical and surgical hospitals in 2007. From the hospital discharge records, I know the total list charges for each visit, as well as the APR-DRG and SOI. I deflate the list charges using the hospital level Medicare cost-to-charge ratios, and use the above formula. The resulting maximum bonuses should be very similar to those used in the Gainsharing Demonstration, as these same records and cost ratios to calculate their bonuses, unless different information was by the hospital.

Realized Bonuses:

The realized bonus is composed of two parts: a performance incentive and an improvement incentive. The performance incentive depends on how much a doctor reduced costs of a particular patient relative to the pool of patients of that type before the program started. The improvement incentive depends on how much a doctor reduced costs of a particular patient relative to their own costs for that type of patient before the program started. For the first year of the program, the weight was 1/3 for the performance incentive and 2/3 for the improvement incentive. Some hospitals changed

these weights to favor the performance incentive over the course of the program.

Realized Bonus Formula for Surgical Patients: The rate year cost is the cost of the index visit, while the 25th, 75th, and 90th percentiles refer to those percentiles of the cost distribution of all patients of a particular type in 2007. The base year cost refers to the costs of the doctor's own patients of the particular type in 2007.

$$\begin{aligned}
 & \underbrace{\frac{1}{3} * MaxBonus * \frac{90^{th} \text{ pctl} - \text{rate yr cost}}{90^{th} \text{ pctl} - 25^{th} \text{ pctl}}}_{\text{Performance Incentive}} + \\
 & \underbrace{\frac{2}{3} * MaxBonus * \frac{\text{base yr cost} - \text{rate yr cost}}{75^{th} \text{ pctl} - 25^{th} \text{ pctl}}}_{\text{Improvement Incentive}}
 \end{aligned} \tag{1.9}$$

Realized Bonus Formula for Medical Patients: The performance incentive is the same, but the improvement incentive is calculated using length of stay rather than costs.

$$\begin{aligned}
 & \underbrace{\frac{1}{3} * MaxBonus * \frac{90^{th} \text{ pctl} - \text{rate yr cost}}{90^{th} \text{ pctl} - 25^{th} \text{ pctl}}}_{\text{Performance Incentive}} + \\
 & \underbrace{\frac{2}{3} * \frac{MaxBonus}{Best \text{ practice LOS}} * (\text{base yr LOS} - \text{rate yr LOS})}_{\text{Improvement Incentive}}
 \end{aligned} \tag{1.10}$$

Due to the fact that the maximum bonus is what matters ex ante, when treatment decisions are being made, I focus on the maximum bonus throughout the paper. In principle I could calculate realized bonuses as well, but I do not for two reasons. First, due to typos and problems with string matching, there is measurement error in my assignment of patients to doctors. This doesn't matter for the creation of maximum bonuses, but if I assigned particularly expensive visit to the wrong doctor in the base year, this would throw off the calculation of the improvement incentive. The

second reason is that the documents detailing the bonus calculation are extremely vague as to what base year cost or base year length of stay is used.

Mathematical Appendix: Model of doctor Decision-making

Doctors make three decisions: whether a patient is admitted, $A \in \{0, 1\}$, whether to admit a patient to a bonus hospital or a regular hospital, $H \in \{0, 1\}$, and how much care to provide, q . Patients vary only by their sickness level $\beta \sim U([0, \bar{\beta}])$. Doctors choose A , H , and q to maximize a weighted average of their profits and the patient's utility from receiving treatment,

$$\begin{aligned} \max_{A,H,q} U(A, H, q; \beta) = & \lambda \underbrace{[aq + \max\{\alpha_0 - \alpha_1 q, 0\} * \mathbb{1}\{H = 1, A = 1\}]}_{\text{Doctor's profits}} \\ & + (1 - \lambda) \underbrace{\left[\beta q + (\gamma q - C) * \mathbb{1}\{A = 1\} - \frac{b}{2} q^2 \right]}_{\text{Patient's utility of treatment}} \end{aligned} \quad (1.11)$$

Doctors' profits in the normal hospital are the amount of services provided, q , multiplied by a reimbursement rate, a . If $A = 1$ and $H = 1$, doctors may also receive a bonus: $\max\{\alpha_0 - \alpha_1 q, 0\}$. The patient's utility function for medical care is concave in q . Sicker patients and admitted patients get more benefit from any treatment, q . Patients also care about admission. Care provided when a patient is admitted is more beneficial (γ), but there is a fixed cost to the patient of admission, C .

Finally, doctors' choices are subject to three restrictions. First, the same number of patients must be admitted at each hospital. Second, Doctors can only admit as many patients as they would admit if there was no bonus. Third, all parameters are in $\mathbb{R}^+$, and $0 < \lambda < 1$.

Pre-Period: Neither Hospital Offers a Bonus

In order to know the capacity constraints that will constrain doctors in the full model, I first solve the model in the absence of the bonus (the “pre-period”). Doctors choose admission, $A \in \{0, 1\}$, and the quantity of care to provide, q . Since both hospitals are identical in the absence of the bonus, and doctors have to admit the same number of patients at each hospital, the hospital choice drops out.

Doctor's choose q and the hospital $A \in \{0, 1\}$ to maximize the utility function:

$$\begin{aligned}
 \max_{A,q} U(A, q; \beta) &= \underbrace{\lambda [aq]}_{\text{profit}} + (1 - \lambda) \underbrace{\left[\beta q + (\gamma q - C) * \mathbb{1}\{A = 1\} - \frac{b}{2} q^2 \right]}_{\text{patient's utility from treatment}} \quad (1.12) \\
 &= \max \left\{ \underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[(\beta + \gamma) q^*(\beta) - C - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_1(\beta) = U(q^*(\beta); \beta, A=1)}, \right. \\
 &\quad \left. \underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[\beta q^*(\beta) - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_0(\beta) = U(q^*(\beta); \beta, A=0)} \right\}
 \end{aligned}$$

Proposition 1: Under some parameter conditions, there exists a β^A such that all patients with $\beta < \beta^A$ are not admitted, and all patients with $\beta \geq \beta^A$ are admitted.

Proof: Need to know the doctor's utility as a function of β .

The value function is:

$$V(\beta) = \max \left\{ \underbrace{\lambda [aq_{(1)}] + (1 - \lambda) \left[(\beta + \gamma) q_{(1)} - C - \frac{b}{2} q_{(1)}^2 \right]}_{V_1(\beta)}, \right. \\ \left. \underbrace{\lambda [aq_{(0)}] + (1 - \lambda) \left[\beta q_{(0)} - \frac{b}{2} q_{(0)}^2 \right]}_{V_0(\beta)} \right\} \quad (1.13)$$

Where from the first order conditions:

$$q_{(1)} = \frac{1}{b} \left[\left(\frac{\lambda}{1 - \lambda} \right) (a_1) + \beta + \gamma \right] \quad (1.14)$$

and

$$q_{(0)} = \frac{1}{b} \left[\left(\frac{\lambda}{1 - \lambda} \right) (a_1) + \beta \right] \quad (1.15)$$

The doctor's utility as a function of β is the upper envelope of $V_1(\beta)$ and $V_0(\beta)$: the utility if all patients are admitted and if all patients are not admitted (see Figure 1.4). Assume the doctor admits all patients with $\beta \in [\beta^A, \bar{\beta}]$, and does not admit patients with $\beta \in [0, \beta^A]$. Now suppose a doctor were to admit a patient with $\beta_1 < \beta^A$. Since $V_1(\beta_1) < V_0(\beta_1)$, a doctor would never choose to admit this patient. Likewise, suppose a doctor were to not admit a patient with $\beta_2 > \beta^A$. Now $V_1(\beta_2) > V_0(\beta_2)$, and again the doctor would be worse off. (See Appendix Figure 1.8). Thus, patients with $\beta \in [\beta^A, \bar{\beta}]$ are all admitted, and the rest are not admitted.

In order to solve the model in the post-period, it is necessary to know β^A . Define β^A such that $U(q_{(0)}, \beta^A) = U(q_{(1)}, \beta^A)$. Therefore, β^A solves:

$$\lambda [aq_{(0)}(\beta^A)] + (1 - \lambda) \left[\beta^A q_{(0)}(\beta^A) - \frac{b}{2} q_{(0)}(\beta^A)^2 \right] = \\ \lambda [aq_{(1)}(\beta^A)] + (1 - \lambda) \left[(\beta + \gamma) q_{(1)}(\beta^A) - C - \frac{b}{2} q_{(1)}(\beta^A)^2 \right] \quad (1.16)$$

$$\Rightarrow \beta^A = \frac{2a\gamma\lambda + 2bC\lambda - 2bC - \gamma^2\lambda + \gamma^2}{2\gamma(\lambda - 1)} \quad (1.17)$$

Post-Period: Hospital 1 Offers a Bonus

Doctors again choose the quantity of care, q , the hospital, $H \in \{0, 1\}$, and admission, $A \in \{0, 1\}$. Now, however, hospital 1 introduces a cost reduction bonus, which is only available for doctors treating admitted patients. The bonus generates a difference between hospitals, and so the hospital choice becomes relevant. In addition, doctors are constrained by the pre-period capacity—they can only admit $\bar{\beta} - \beta^A$ patients, and they must distribute the admitted patients evenly across hospitals.

Doctors choose q , $H \in \{0, 1\}$, and $A \in \{0, 1\}$ to maximize the utility function

$$\begin{aligned} \max_{A, H, q} U(A, H, q; \beta) &= \underbrace{\lambda [aq + \max\{\alpha_0 - \alpha_1 q, 0\} * \mathbb{1}\{H = 1, A = 1\}]}_{\text{profit}} \\ &\quad + (1 - \lambda) \underbrace{\left[\beta q + (\gamma q - C) * \mathbb{1}\{A = 1\} - \frac{b}{2} q^2 \right]}_{\text{patient's utility from treatment}} \\ &= \max \left\{ \begin{aligned} &\lambda [aq^*(\beta) + \alpha_0 - \alpha_1 q^*(\beta)] + \\ &\underbrace{(1 - \lambda) \left[(\beta + \gamma) q^*(\beta) - C - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_2(\beta) = U(q^*(\beta); \beta, H=1, A=1)}, \\ &\underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[(\beta + \gamma) q^*(\beta) - C - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_1(\beta) = U(q^*(\beta); \beta, H=0, A=1)}, \\ &\underbrace{\lambda [aq^*(\beta)] + (1 - \lambda) \left[\beta q^*(\beta) - \frac{b}{2} q^*(\beta)^2 \right]}_{\mathbf{V}_0(\beta) = U(q^*(\beta); \beta, A=0)} \end{aligned} \right\} \end{aligned} \quad (1.18)$$

Subject to the capacity constraint: a maximum of β' patients can be admitted at each hospital,

where $\beta' = \frac{\bar{\beta} - \beta^A}{2}$

Proposition 2: Under some parameter restrictions, there exists a $\tilde{\beta}$ such that patients with $\beta \in [0, \tilde{\beta}]$ are admitted at the bonus hospital, patients with $\beta \in [\tilde{\beta}, \tilde{\beta} + \beta^A]$ are not admitted, and the remaining patients with $\beta \in [\tilde{\beta} + \beta^A, \bar{\beta}]$ are admitted at either the bonus or non-bonus hospital.

Proof: Need to know the doctor's utility as a function of β .

The value function is

$$V(\beta) = \max \left\{ \underbrace{\lambda [aq_{(1)} + \alpha_0 - \alpha_1 q_{(1)}] + (1 - \lambda) \left[(\beta + \gamma) q_{(1)} - C - \frac{b}{2} q_{(1)}^2 \right]}_{V_2(\beta)}, \right. \quad (1.19)$$

$$\underbrace{\lambda [aq_{(2)}] + (1 - \lambda) \left[(\beta + \gamma) q_{(2)} - C - \frac{b}{2} q_{(2)}^2 \right]}_{V_1(\beta)},$$

$$\left. \underbrace{\lambda [aq_{(0)}] + (1 - \lambda) \left[\beta q_{(0)} - \frac{b}{2} q_{(0)}^2 \right]}_{V_0(\beta)} \right\}$$

subject to the capacity constraint; only β' patients can be admitted to each hospital.

i. If the doctor chooses q under the first term, it must satisfy the following FOC:

$$q_{(2)} = \frac{1}{b} \left[\left(\frac{\lambda}{1 - \lambda} \right) (a - \alpha_1) + \beta + \gamma \right] \quad (1.20)$$

ii. If the doctor chooses q under the second term, it must satisfy the following FOC:

$$q_{(1)} = \frac{1}{b} \left[\left(\frac{\lambda}{1 - \lambda} \right) (a) + \beta + \gamma \right] \quad (1.21)$$

iii. If the doctor chooses q under the third term, it must satisfy the following FOC:

$$q_{(0)} = \frac{1}{b} \left[\left(\frac{\lambda}{1-\lambda} \right) (a) + \beta \right] \quad (1.22)$$

Under certain conditions, the value function of the doctor is composed of three segments of the three parts of the value function, which maximize total utility (see Figure 1.5). Assume doctors decide which patients to admit and where to admit them by dividing their patients into three segments of β . They then admit the low β patients to the bonus hospital, do not admit the middle β s, and admit the highest β s to either hospital (randomizing over hospital such that they admit β' patients at both hospitals). Define the cut points as $\tilde{\beta}$ and $\tilde{\beta} + \beta^A$. There is no patient β_2 with $\tilde{\beta} \leq \beta_2 \leq \tilde{\beta} + \beta^A$ where the doctor would prefer to admit β_2 if it meant giving up admission for any patient $\beta_1 < \tilde{\beta}$ or $\beta_3 > \tilde{\beta} + \beta^A$; the doctor would be strictly worse off. This situation is depicted in Appendix Figure 1.9.

The $\tilde{\beta}'$ that partitions the range of β into these three groups solves $U(q_{(2)}, \tilde{\beta}') - U(q_{(0)}, \tilde{\beta}') = U(q_{(1)}, \tilde{\beta}' + \beta^A) - U(q_{(0)}, \tilde{\beta}' + \beta^A)$:

$$\begin{aligned} \tilde{\beta}' \text{ solves: } & \left(\lambda [aq_{(2)} + \alpha_0 - \alpha_1 q_{(2)}] + (1-\lambda) \left[(\tilde{\beta} + \gamma) q_{(2)} - C - \frac{b}{2} q_{(2)}^2 \right] \right) \\ & - \left(\lambda [aq_{(0)'}] + (1-\lambda) \left[\tilde{\beta} q_{(0)'} - \frac{b}{2} q_{(0)'}^2 \right] \right) \\ & = \left(\lambda [aq_{(1)}] + (1-\lambda) \left[(\tilde{\beta} + \beta^A + \gamma) q_{(1)} - C - \frac{b}{2} q_{(1)}^2 \right] \right) \\ & - \left(\lambda [a_1 q_{(0)''}] + (1-\lambda) \left[(\tilde{\beta} + \beta^A) q_{(0)''} - \frac{b_2}{2} q_{(0)''}^2 \right] \right) \end{aligned} \quad (1.23)$$

$$\text{where: } q_{(2)} = \frac{1}{b} \left[\left(\frac{\lambda}{1-\lambda} \right) (a - \alpha_1) + \tilde{\beta} + \gamma \right] \quad (1.24)$$

$$q_{(0)'} = \frac{1}{b_2} \left[\left(\frac{\lambda}{1-\lambda} \right) (a) + \tilde{\beta} \right] \quad (1.25)$$

$$q_{(1)} = \frac{1}{b} \left[\left(\frac{\lambda}{1-\lambda} \right) (a) + \tilde{\beta} + \beta^A + \gamma \right] \quad (1.26)$$

$$q_{(0)''} = \frac{1}{b} \left[\left(\frac{\lambda}{1-\lambda} \right) (a) + \tilde{\beta} + \beta^A \right] \quad (1.27)$$

$$\beta^A = \frac{2a\gamma\lambda + 2bC\lambda - 2bC - \gamma^2\lambda + \gamma^2}{2\gamma(\lambda - 1)} \quad (1.28)$$

$$\tilde{\beta}' = \frac{\lambda(\lambda\alpha_1(2a-2\gamma-\alpha_1)+2b\alpha_0(\lambda-1)+2\gamma\alpha_1)+(\lambda-1)(2a\gamma\lambda+2bC\lambda-2bC-\gamma^2\lambda+\gamma^2)}{2\lambda\alpha_1(\lambda-1)} \quad (1.29)$$

However, because of the capacity constraint, $\tilde{\beta} = \min \left\{ \tilde{\beta}', \frac{\tilde{\beta} - \beta^A}{2} \right\}$.

Proposition 3: The direction of the change in q conditional on β from the pre- to the post-period for bonus-generating patients ($\beta \in [0, \tilde{\beta}]$) is ambiguous.

Proof: If a patient would be admitted even without the bonus (in the pre-period), the introduction of the bonuses is associated with a lower q . For patients who are not admitted in the absence of the bonuses, however, the relevant comparison is between the q chosen under the bonus scheme ($q_{(2)}$), and the q chosen when a patient is not admitted ($q_{(0)}$). From the first order conditions of the doctor's value function, the optimal q when a patient is admitted at the bonus hospital is $q_{(2)} = \frac{1}{b} \left[\left(\frac{\lambda}{1-\lambda} \right) (a - \alpha_1) + \beta + \gamma \right]$, and the optimal q when a patient is not admitted is $q_{(0)} = \frac{1}{b} \left[\left(\frac{\lambda}{1-\lambda} \right) (a) + \beta \right]$. Whether the quantity of care provided for the bonus generating patients is higher or lower than the counterfactual of neither hospital offering a bonus is determined by the relative size of γ and α_1 . If $\frac{1}{b} \left[\gamma - \left(\frac{\lambda}{1-\lambda} \right) (\alpha_1) \right] < 0$, the quantity of care provided for patients with $\beta \in [0, \tilde{\beta}]$ is less when hospital 1 implements the bonus scheme than when neither hospital implements the bonus. On the other hand, if $\frac{1}{b} \left[\gamma - \left(\frac{\lambda}{1-\lambda} \right) (\alpha_1) \right] > 0$, the quantity of care provided for patients with $\beta \in [0, \tilde{\beta}]$ is greater

when hospital 1 implements the bonus scheme than when neither hospital implements the bonus.

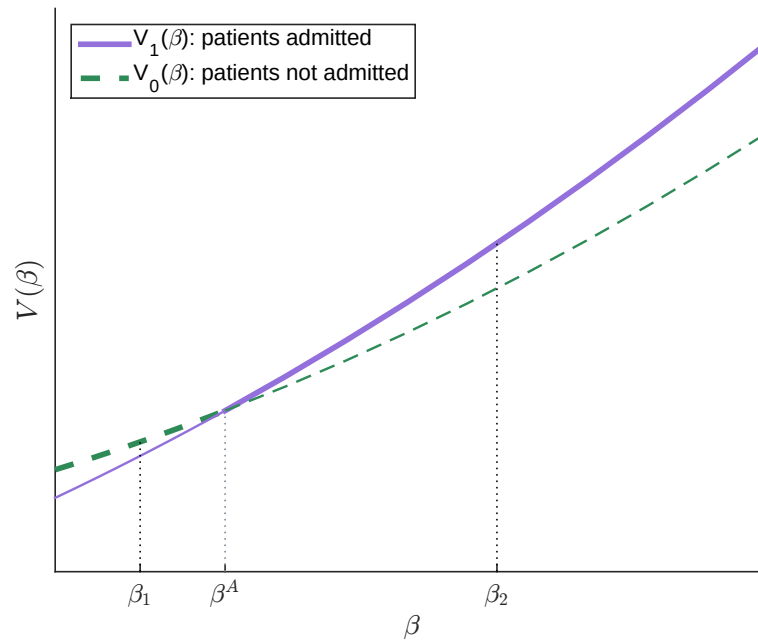
Appendix Figure 1.10 shows both cases: 1.10A demonstrates the case where the quantity of care provided for the bonus generating patients is less under the bonus program than the counterfactual of no bonuses; 1.10B shows the opposite.

Parameter conditions

The above interior solution exists as long as three sets of parameter restrictions hold. First, $V_0(\beta)$ and $V_1(\beta)$ cross; in the absence of the bonus, some patients are admitted and some patients are not admitted. Second, the bonuses are large enough to matter; the bonuses induce the doctor to admit the healthiest patient over the “healthiest” of the sick patients they formerly admitted. The second condition holds as long as $U(q_{(2)}, 0) - U(q_{(0)}, 0) > U(q_{(1)}, \beta^A) - U(q_{(0)}, \beta^A)$. Finally, the doctors always want to admit the sickest patients: $U(q_{(2)}, \tilde{\beta}) - U(q_{(0)}, \tilde{\beta}) < U(q_{(1)}, \tilde{\beta}) - U(q_{(0)}, \tilde{\beta})$.

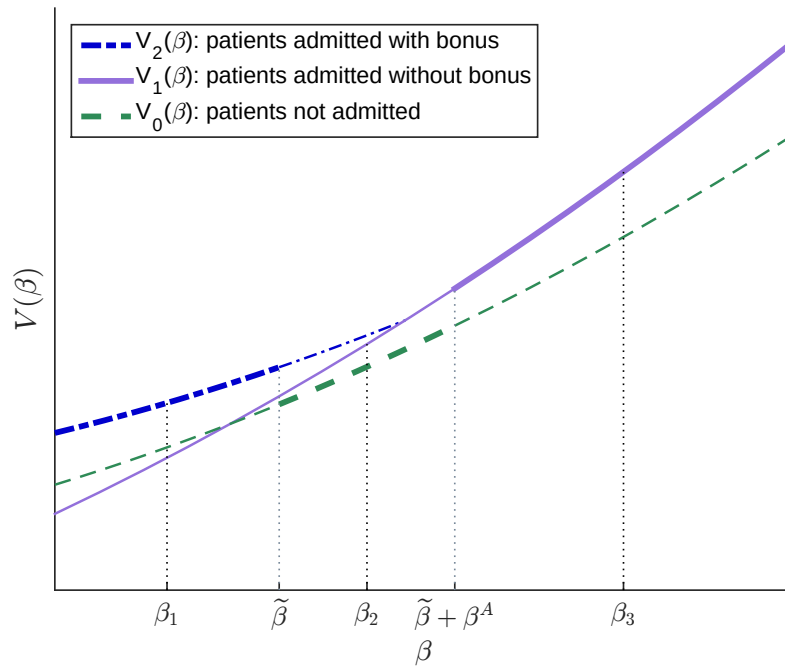
1.9.1 Appendix Figures

Figure 1.8: Doctor's Utility as a Function of β : without Bonuses



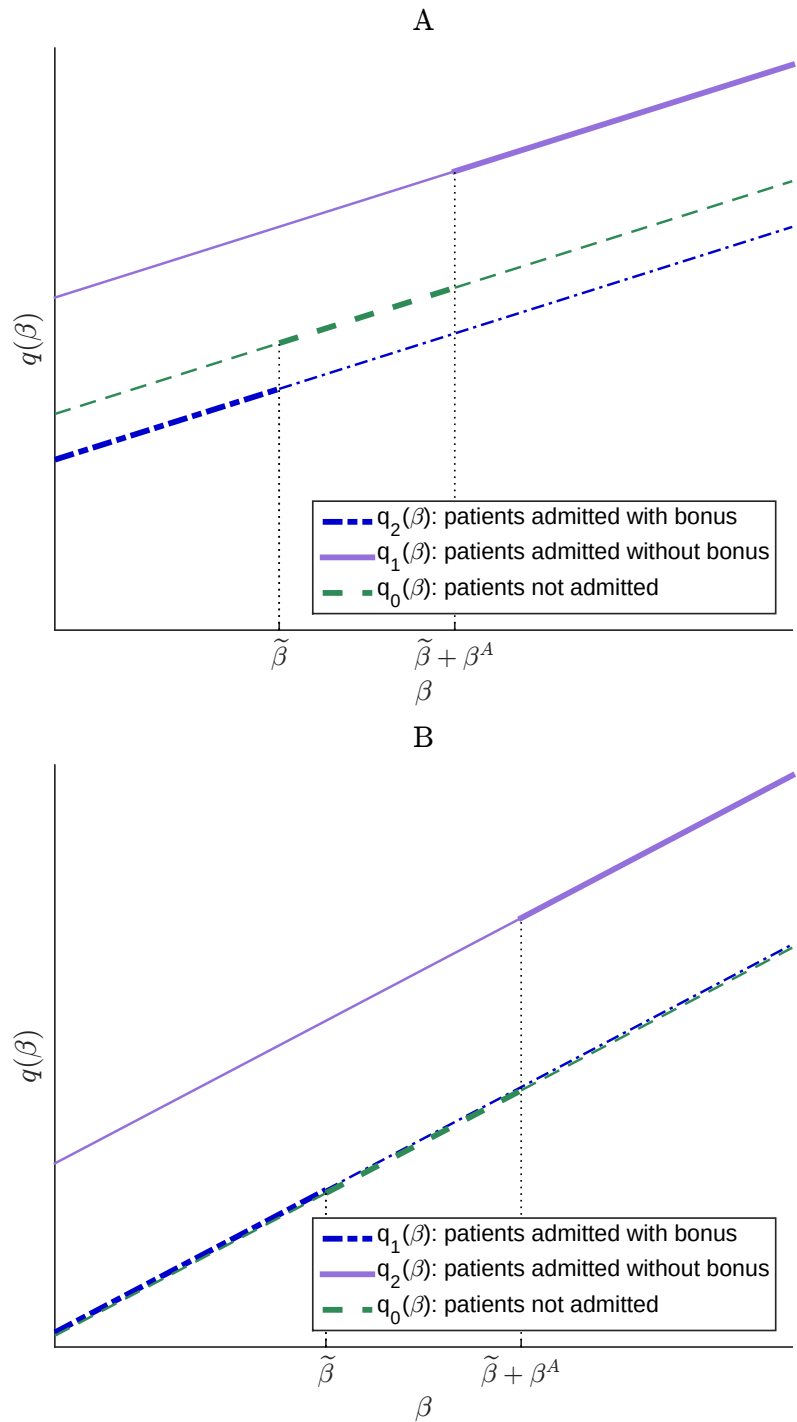
Notes: The bold line sections show the optimal decision rule as a function of β .

Figure 1.9: Doctor's Utility as a Function of β : with Bonuses



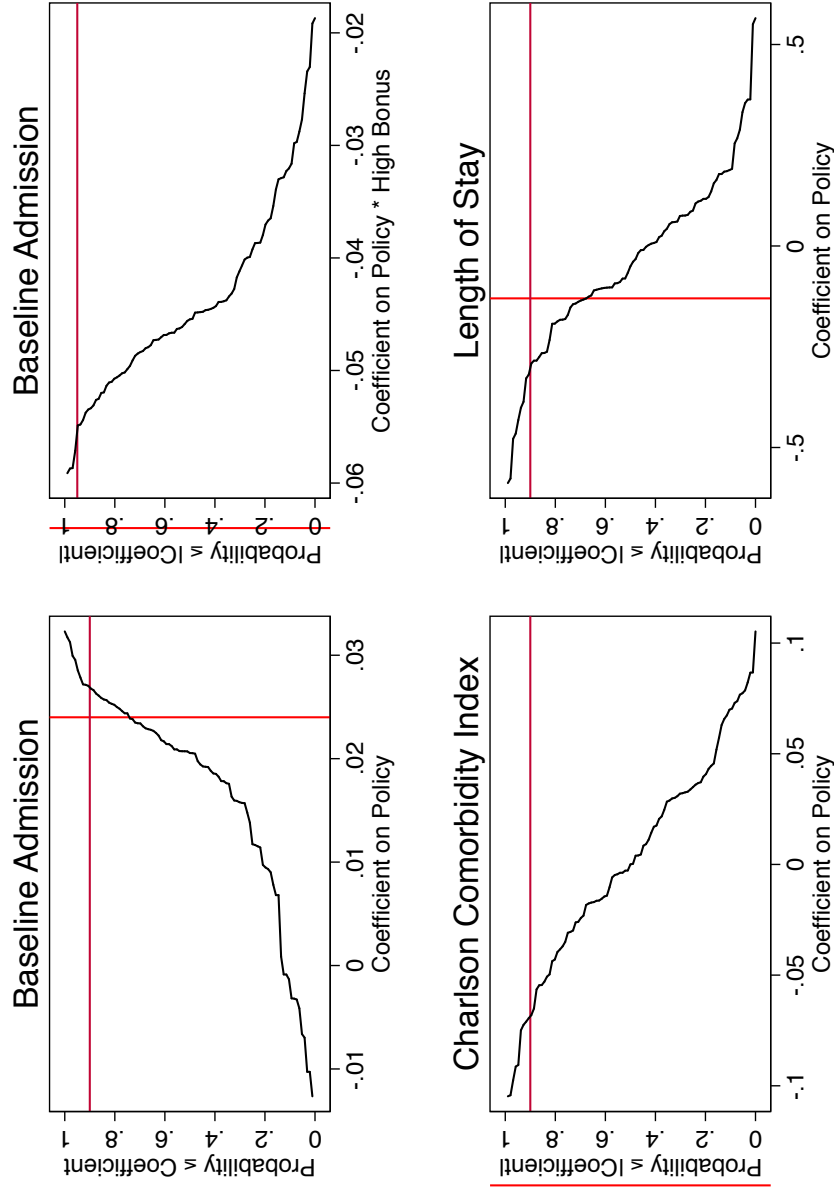
Notes: The bold line sections show the optimal decision rule as a function of β .

Figure 1.10: Optimal Quantity of Care as a Function of β : with Bonuses



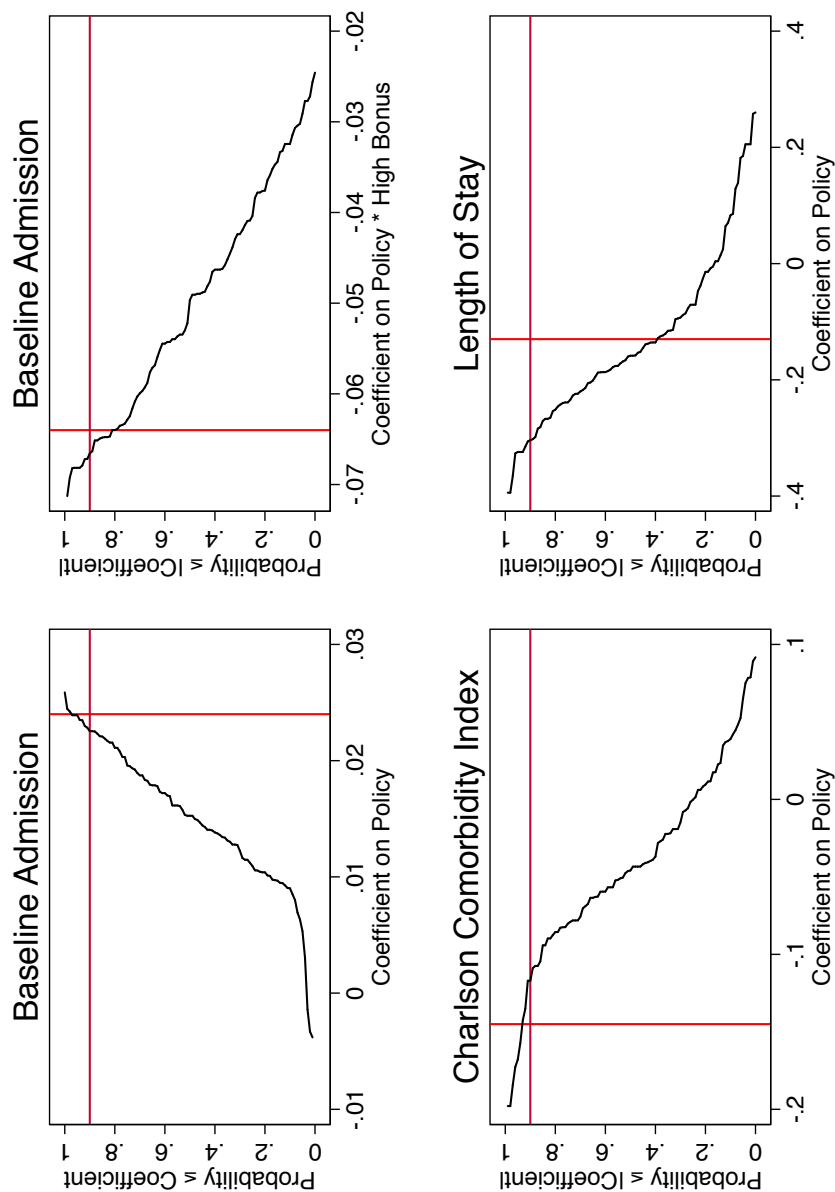
Notes: The bold line sections show the quantity of care provided along the optimal decision rule. Figure 1.10A shows the optimal quantity of care under one set of parameters; Figure 1.10B shows the optimal quantity of care under another set of parameters.

Figure 1.11: Placebo Test: Randomize Program Participation



Notes: Based on 100 repetitions, with the participation into the the Gainsharing Demonstration and the BPCI Model 1 program randomly assigned (holding fixed the number of hospitals participating in each at the true value). The vertical bar denotes the estimate in the real data, and the horizontal bar is placed at the 90th percentile.

Figure 1.12: Placebo Test: Randomize Dates



Notes: Based on 100 repetitions, with the starting dates of the Gainsharing Demonstration and the BPCI Model 1 program randomly assigned (two random dates are chosen; the first is assigned to the Gainsharing Demonstration, and the second to the BPCI Model 1 program). The vertical bar denotes the estimate in the real data, and the horizontal bar is placed at the 90th percentile.

1.9.2 Appendix Tables

Table 1.11: Top 15 Diagnoses for Medical and Surgical Patients
(CCS Diagnosis Codes)

Medical Patients			
Primary Diagnosis	Frequency	Percent	Cumulative
Diseases of the heart	128,993	16.06	16.06
Diseases of the urinary system	54,048	6.73	22.7
Upper gastrointestinal disorders	37,511	4.67	27.4
Complications	30,919	3.85	31.3
Symptoms; signs; and ill-defined condit	28,173	3.51	34.8
Benign neoplasms	28,106	3.5	38.3
Lower gastrointestinal disorders	27,110	3.38	41.6
Spondylosis; intervertebral disc disord	26,493	3.3	44.9
Cerebrovascular disease	23,615	2.94	47.9
Respiratory infections	20,821	2.59	50.5
Chronic obstructive pulmonary disease a	16,985	2.11	52.6
Anemia	16,504	2.05	54.6
Eye disorders	15,052	1.87	56.5
Diseases of arteries; arterioles; and c	14,117	1.76	58.3
Other nervous system disorders	13,446	1.67	60
Surgical Patients			
Primary Diagnosis	Frequency	Percent	Cumulative
Diseases of the heart	84,362	14.52	14.5
Eye disorders	65,469	11.27	25.8
Non-traumatic joint disorders	51,587	8.88	34.6
Complications	34,379	5.92	40.6
Abdominal hernia	24,618	4.24	44.8
Fractures	24,449	4.21	49
Spondylosis; intervertebral disc disord	22,213	3.82	52.8
Diseases of female genital organs	20,387	3.51	56.3
Diseases of arteries; arterioles; and c	20,056	3.45	59.8
Diseases of the urinary system	16,338	2.81	62.6
Cancer of urinary organs	14,540	2.5	65.1
Biliary tract disease	14,131	2.43	67.5
Cancer of breast	13,816	2.38	69.9
Cerebrovascular disease	11,506	1.98	71.9
Diseases of male genital organs	10,823	1.86	73.8

Table 1.12: Participating Hospitals

Gainsharing Demonstration	Gainsharing Extension
AtlantiCare Regional Medical Center	CentraState Healthcare System
Overlook Hospital	Hunterdon Medical Center
Holy Name Hospital	Jersey Shore University Med. Center
Jersey Shore University Medical Center	JFK Medical Center
Hunterdon Medical Center	Monmouth Medical Center
Monmouth Medical Center	Overlook Hospital
St. Francis Medical Center	St. Francis Medical Center
Our Lady of Lourdes Medical Center	The Valley Hospital
The Valley Hospital	
Somerset Medical Center	
JFK Medical Center	
CentraState Healthcare System	
BPCI Model 1 Program	
Capital Health Medical Center-Hopewell	RWJ University Hospital- Rahway
Capital Health Regional Medical Center	Saint Clare's Hospital- Denville
CentraState Medical Center	Saint Clare's Hospital- Dover
Cooper Hospital/University Medical Center	Saint Michael's Medical Center
Deborah Heart and Lung Center	Saint Peter's University Hospital
Hunterdon Medical Center	South Jersey Healthcare- Elmer
Jersey Shore University Medical Center	South Jersey Healthcare- RMC
JFK Medical Center	St. Joseph's RMC
Morristown Medical Center	St. Mary's Hospital Passaic
Overlook Medical Center	The Valley Hospital
Robert Wood Johnson (RWJ) University Hosp.	Underwood-Memorial Hospital
RWJ University Hospital- Hamilton	University Medical Center of Princeton

Notes: RMC stands for regional medical center.

Table 1.13: Effect of Program on Ex-Ante Patient Health: Medicare ER Patients

	Medical Patients		Surgical Patients	
	CCI Past	Tot. Chronic	CCI Past	Tot. Chronic
policy	-0.04 (0.02)	-0.02 (0.04)	-0.05 (0.04)	-0.06 (0.06)
Doctor FEs	✓	✓	✓	✓
Mean	2.562	5.159	1.918	4.504
Clusters	78	78	76	76
<i>N</i>	1,592,186	1,592,186	95,542	95,542

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. CCI stands for Charlson Co-morbidity Index, which is calculated based on information in previous visits. Tot. Chronic refers to the number of body systems affected by chronic conditions. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.14: Effect of Program on Baseline Admission Probability: Medicare ER Patients

	Medical Patients		Surgical Patients	
	(1)	(2)	(3)	(4)
	Baseline Adm.	Baseline Adm.	Baseline Adm.	Baseline Adm.
policy	0.003 (0.004)	0.002 (0.004)	0.000 (0.002)	-0.001 (0.002)
high bonus		0.317*** (0.011)		0.497*** (0.053)
policy * high bonus		0.004 (0.003)		0.001 (0.002)
Doctor FEs	✓	✓	✓	✓
Mean	0.752	0.752	0.969	0.969
Clusters	78	78	76	76
<i>N</i>	1,585,753	1,585,753	69,742	69,742

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.15: Effect of Program on Costs and Procedure Use: Medicare ER Patients

Panel A: Medical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Any Imaging	Diag. Ultra	Total Costs
policy	-0.026 (0.084)	-0.003 (0.005)	-0.003 (0.003)	0.004 (0.007)	0.009* (0.005)	471.649 (387.532)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	5.211	0.0797	0.0306	0.135	0.0364	8,300.5
Clusters	78	78	78	78	78	65
<i>N</i>	1,592,186	1,592,186	1,592,186	1,592,186	1,592,186	1,341,584

Panel B: Surgical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Any Imaging	Diag. Ultra	Total Costs
policy	-0.249 (0.236)	-0.003 (0.006)	-0.002 (0.005)	0.002 (0.010)	0.007 (0.007)	249.986 (584.883)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	7.655	0.0746	0.0268	0.202	0.0485	17,000.0
Clusters	76	76	76	76	76	65
<i>N</i>	95,542	95,542	95,542	95,542	95,542	82,959

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race, and the variables measuring underlying health from Table 1.5. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.16: Effect of Program on Baseline Admission Probability: Near Medicare Patient

	Medical Patients		Surgical Patients	
	(1)	(2)	(3)	(4)
	Baseline Adm.	Baseline Adm.	Baseline Adm.	Baseline Adm.
policy	0.000 (0.003)	0.004 (0.004)	-0.004 (0.005)	-0.001 (0.004)
high bonus		0.028 (0.028)		0.609*** (0.020)
policy * high bonus		-0.011** (0.004)		-0.006** (0.002)
Doctor FEs	✓	✓	✓	✓
Mean	0.699	0.699	0.845	0.845
Clusters	79	79	78	78
<i>N</i>	183,259	183,259	247,488	247,488

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.17: Effect of Program on Ex-Ante Patient Health: Near Medicare Patients

	Medical Patients		Surgical Patients	
	CCI Past	Tot. Chronic	CCI Past	Tot. Chronic
policy	-0.06 (0.04)	-0.05 (0.03)	-0.02 (0.02)	-0.01 (0.03)
Doctor FEs	✓	✓	✓	✓
Mean	1.866	4.162	0.885	3.531
Clusters	79	79	78	78
<i>N</i>	187,054	187,054	256,141	256,141

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. CCI stands for Charlson Co-morbidity Index, which is calculated based on information in previous visits. Tot. Chronic refers to the number of body systems affected by chronic conditions. Standard errors clustered at the hospital level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.18: Effect of Program Costs and Procedure Use: Near Medicare Patients

Panel A: Medical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Any Imaging	Diag. Ultra	Total Costs
	b/se	b/se	b/se	b/se	b/se	b/se
policy	-0.020 (0.162)	-0.000 (0.004)	-0.001 (0.003)	0.000 (0.007)	0.003 (0.004)	612.815 (527.039)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	6.181	0.0344	0.0190	0.115	0.0422	10,459.2
Clusters	79	79	79	79	79	66
<i>N</i>	187,054	187,054	187,054	187,054	187,054	173,591

Panel B: Surgical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Any Imaging	Diag. Ultra	Total Costs
	b/se	b/se	b/se	b/se	b/se	b/se
policy	0.056 (0.103)	-0.003 (0.003)	-0.000 (0.001)	0.008 (0.009)	-0.002 (0.004)	1,032.178** (458.950)
Doctor FEs	✓	✓	✓	✓	✓	✓
Mean	5.052	0.0180	0.00716	0.135	0.0458	16,148.8
Clusters	78	78	78	78	78	66
<i>N</i>	256,141	256,141	256,141	256,141	256,141	230,910

Notes: Quarter, doctor, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race, and the variables measuring underlying health from Table 1.5. Standard errors clustered at the hospital level. *p < 0.1, ** p < 0.05, *** p < 0.01

Table 1.19: Effect of Program on Baseline Admission Probability: Single Hospital Doctors

	Medical Patients		Surgical Patients	
	(1)	(2)	(3)	(4)
	Baseline Adm.	Baseline Adm.	Baseline Adm.	Baseline Adm.
policy	0.001 (0.009)	0.031*** (0.008)	-0.005 (0.007)	0.010 (0.009)
high bonus		-0.080** (0.030)		-0.058*** (0.007)
policy * high bonus		-0.077*** (0.007)		-0.024** (0.011)
Mean	0.911	0.911	0.947	0.947
Clusters	70	70	71	71
<i>N</i>	173,619	173,619	120,363	120,363

Notes: Quarter, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. Standard errors clustered at the hospital level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.20: Effect of Program on Ex-Ante Patient Health: Single Hospital Doctors

	Medical Patients		Surgical Patients	
	CCI Past	Tot. Chronic	CCI Past	Tot. Chronic
policy	-0.05 (0.06)	0.02 (0.07)	-0.02 (0.04)	0.02 (0.06)
Mean	2.653	4.135	1.603	2.737
Clusters	70	70	71	71
<i>N</i>	175,190	175,190	129,225	129,225

Notes: Quarter, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race. CCI stands for Charlson Co-morbidity Index, which is calculated based on information in previous visits. Tot. Chronic refers to the number of body systems affected by chronic conditions. Standard errors clustered at the hospital level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.21: Effect of Program Costs and Procedure Use: Single Hospital Doctors

Panel A: Medical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Diag. Ultra	Any Imaging	Total Costs
policy	0.091 (0.304)	0.004 (0.003)	-0.001 (0.004)	-0.001 (0.004)	0.001 (0.007)	619.496 (660.191)
Mean	7.174	0.0194	0.0130	0.0334	0.0890	11118.3
Clusters	70	70	70	70	70	57
<i>N</i>	175,190	175,190	175,190	175,190	175,190	169,118

Panel B: Surgical Patients

	(1)	(2)	(3)	(4)	(5)	(6)
	Length of Stay	CT Scan	MRI	Diag. Ultra	Any Imaging	Total Costs
policy	0.127 (0.124)	-0.002 (0.004)	-0.000 (0.001)	0.011 (0.010)	0.020* (0.011)	1,730.635** (738.486)
Mean	6.871	0.0199	0.00710	0.0720	0.173	19727.8
Clusters	71	71	71	71	71	58
<i>N</i>	129,225	129,225	129,225	129,225	129,225	120,393

Notes: Quarter, hospital, and diagnosis by severity of illness fixed effects also included, as well as dummies for age categories, sex, and race, and the variables measuring underlying health from Table 1.5. Standard errors clustered at the hospital level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2

Just What the Nurse Practitioner Ordered: Independent Prescriptive Authority and Population Mental Health*

2.1 Introduction

Limited access to mental health care services in the United States is a major public health concern. While one in five Americans suffers from a mental illness, nearly one third of the US currently lives in areas that are underserved by mental health care providers (Bureau of Health Workforce, Bureau of Health Workforce (2016)). As mental health problems tend to develop early in life and persist over the lifecycle, the costs of not receiving treatment can be devastating. In addition to direct medical costs, untreated mental illnesses are associated with lower human capital accumulation, worse labor market participation and performance, and greater criminal activity Currie and Stabile (2006); Ettner et al. (1997); Greenberg and Rosenheck (2008). Recent evidence demonstrates that even increases in all-cause mortality for some groups are being driven by mental-health-related deaths, turning the policy focus to the challenging question of what can be done to improve population mental health Case and Deaton (2015).

In this paper, we examine whether allowing nurse practitioners (NPs)—registered nurses with

*Materials from this chapter were presented at the Twenty-First Annual Meetings of the Society of Labor Economists (2016). We thank Janet Currie, David Krol, Ilyana Kuziemko, Beth Rom-Rymer, and seminar participants at Princeton University for useful comments and suggestions. Generous financial support from the Center for Health and Wellbeing at Princeton University is gratefully acknowledged.

an advanced degree in nursing—to prescribe medication without physician supervision or collaboration is associated with improved mental health outcomes. In doing so, we make two key contributions. First, we show that states that broaden prescriptive authority see both improvements in self-reported mental health and decreases in the number of mental-health-related deaths, including suicides. These benefits are most pronounced in low socioeconomic status populations, and in populations residing in areas with fewer psychiatrists per capita. Second, we show that independent prescriptive authority for NPs is not associated with a crowding out of the services provided by physicians. Together, our results suggest that granting independent prescriptive authority to NPs is an important policy tool which may be used to improve health outcomes for populations with limited access to care.

Broadly speaking, there are two types of treatment for mental illness: psychotherapy and psychotropic medication. A complementarity between the two has been well documented, and in most cases it is recommended that a patient receive a combination of both treatments (SAMHSA, Substance Abuse and Mental Health Services Administration (2015)). Despite this ideal of psychotherapy in conjunction with psychotropic medication, it is often much easier to find consistent access to therapy than to medication. While all mental health professionals can offer some degree of counseling services, traditionally only psychiatrists and other medical doctors have the legislative authority to prescribe medications.

The observed imbalance between access to psychotherapy and access to psychotropic medications is greatest among already disadvantaged populations. Psychiatrists are more likely to locate in urban and suburban areas, leaving populations in rural areas with the most limited access to psychotropic treatment Hartley et al. (2004). Even in areas where there are a sufficient number of psychiatrists, access to psychotropic medications may still be limited for certain populations. Psychiatrists are less likely than all other physician specialties to accept insurance, with differences in acceptance rates being greatest among forms of public insurance Bishop et al. (2014). Conversely, NPs are more likely than physicians to locate in rural and inner-city locations and to accept public insurance Buerhaus et al. (2015). Granting independent prescriptive authority to NPs thus has the

potential both to address physician shortages and to extend care to disadvantaged populations.

We find that extending prescriptive authority to NPs is associated with improvements in both self-reported mental health and mental-health-related mortality. Our identification strategy exploits state-level law changes that dictate the level of independence that NPs and other Advanced Practice Registered Nurses (APRNs) have to diagnose, treat, and prescribe medications from 1990 to 2014. We use within-state variation in this scope of practice legislation to examine how both population mental health and prescribing behaviors of physicians change when NPs are granted independent prescriptive authority. On average, granting NPs independent prescriptive authority is associated with a significant reduction in the number of days spent in poor mental health (0.14 days, or 4% of the mean).

Improvements in mental health outcomes are more pronounced and more precisely estimated both in areas that are underserved by psychiatrists and among already disadvantaged populations. In particular, areas that are underserved by psychiatrists see improvements in self-reported mental health that are at least twice as large as that of an average respondent, as well as significant reductions in mental-health-related mortality. For days in poor mental health, populations with low levels of education see improvements that are even greater: the average low-education resident in a state with fewer psychiatrists per capita sees improvements that are 135% greater than the average resident in the full sample.

Despite a burgeoning literature demonstrating that NPs can safely and efficiently provide a variety of services, including an endorsement of the skills of NPs by the Institute of Medicine, efforts to extend prescriptive authority beyond physicians are controversial. Opponents worry that allowing NPs to prescribe medication would put patients in danger since NPs receive fewer years of training, are held to different legal standards, and go through a different process of licensing than medical doctors. Critics further note that extending prescriptive authority beyond physicians need not expand overall use of pharmacological treatment, as the prescriptions written by non-physician providers may simply crowd out the prescriptions previously written by physicians. The American Medical Association (AMA), a national professional organization representing physicians and

medical students in the US, has been particularly vocal in opposing the expansion of state-level scope of practice legislation.

Using detailed prescription data on the number of prescriptions written by physicians for both antipsychotics and antidepressants, we examine whether extending prescriptive authority to NPs crowds out the services provided by physicians. Again exploiting time-series variation in scope of practice legislation, we find no evidence of crowd out. In fact, in all 18 states that granted NPs independent prescriptive authority between 1990 and 2014, there is no statistically significant difference between the number of prescriptions written by physicians before and after the change in legislation.

Our results indicate that liberalizing scope of practice legislation for non-physician providers can help mitigate the impacts of limited access to physician-provided health care. In particular, states which are underserved by psychiatrists can grant independent prescriptive authority to NPs to decrease the prevalence of negative mental health outcomes. The potential for such legislative action remains large: as of January, 2015, only 24 states and the District of Columbia had granted independent prescriptive authority to NPs. Noticeably, no state in the south has yet to allow NPs to independently prescribe. Given the limited access to psychiatrists in these states, as well as their record of poor mental health outcomes, our findings are of particular importance for the provision of mental health care services in the southern US.

This paper proceeds as follows. In Section 2.2, we introduce the data. In Section 2.3, we examine how self-reported mental health and mental-health-related mortality change when independent prescriptive authority is extended to NPs. Using detailed prescription data, in Section 2.4 we examine how the number of prescriptions for antidepressants and antipsychotics written by physicians change when NPs are granted independent prescriptive authority. Section 2.5 concludes.

2.2 Data

We use information from seven sources to document how extending prescriptive authority to NPs affects population mental health. In particular, we combine a new dataset detailing the prescriptive authority for NPs from 1990 to 2014 with both mental health outcomes from the US mortality files and the Behavioral Risk Factor Surveillance System survey and prescription data from IMS Health’s Xponent database. These data are supplemented with information on the provision of local medical resources from the Area Resource Files, the American Community Survey, and the US Census. Each dataset is described in detail below.

2.2.1 Independent Prescriptive Authority

Our first dataset documents whether NPs had the legislative authority to independently prescribe medication in each month from 1990 to 2014 in each state and the District of Columbia. This dataset was constructed by the authors and combines information from the *The Nurse Practitioner’s* yearly “Annual Legislative Update,” correspondences with state nursing boards, and readings of primary source legislation. While the language of scope of practice legislation is particular to each state, we define independent prescriptive authority as the ability to prescribe medication without physician collaboration or supervision. More information on the construction of this dataset can be found in the Data Appendix.

As of the start of our dataset on January 1st, 1990, six states and the District of Columbia had already granted NPs statutory authority to independently prescribe medication. Between 1990 and 2014, 18 states changed their scope of practice legislation to allow NPs to prescribe without physician involvement. The geographic and temporal variation in scope of practice legislation is displayed in Figure 2.1.

While it is difficult to say why states decide to grant NPs independent prescriptive authority, we believe that the timing of state-level changes in scope of practice legislation are exogenous to population mental health. Anecdotal evidence suggests that changes in scope of practice legislation

are driven by idiosyncrasies of local politics. We find no evidence that the law changes are driven by measures of local economic conditions or the availability of mental health care (see Appendix Table 2.6). Therefore, we believe that these law changes are exogenous to baseline levels and trends of mental health.

2.2.2 Mental Health Outcomes

We identify the impact of independent prescriptive authority for NPs on population mental health using two complementary outcomes: self-reported days in “poor mental health” and mental-health-related mortality. For each outcome, we consider how extending prescriptive authority to NPs impacts both the local population as a whole and disadvantaged subpopulations. As psychiatrists are unlikely to accept public insurance, populations with low levels of income and education may not be able to make use of psychiatrist-provided care Bishop et al. (2014). NPs, on the other hand, are more likely than physicians to treat Medicaid recipients and other vulnerable populations, and thus these populations may benefit more than others from changes in scope of practice legislation Buerhaus et al. (2015). We therefore look separately at low socioeconomic (SES) populations to account for the fact that even when there are sufficient psychiatrists within an area, low SES residents may find it more difficult to access existing care.

Mental-Health-Related Mortality

Our first measure of mental health is mental-health-related mortality at the county level from the US Mortality Files. Here, we consider both suicides and a broader measure of “mental-health-related deaths,” which combines suicides, deaths of unknown intent, and accidental death categories that are closely related to mental health: those involving firearms, trains, drownings, and poisonings Björkenstam et al. (2014); Hilkevitch (2005); Rockett et al. (2006).

Our preferred outcome is the broad measure of mental-health-related mortality for two reasons. First, geographic variation in reported suicides may reflect both systematic differences in true suicides as well as systematic differences in cause-of-death reporting. When someone dies from

an overdose of oxycodone, for example, the local coroner decides whether to label the death as a suicide or as an accidental poisoning. Our broad measure of mental-health-related deaths captures both causes of death, whereas “suicides” only captures the former. Second, drug and alcohol addiction is an increasingly important category of mental health illness, and one which would not be fully captured in suicides Case and Deaton (2015). Therefore, we believe that our broad measure of mental-health-related mortality is more indicative of true differences in mortality caused by poor mental health than suicides alone.

The mortality files contain some demographic information for the deceased individual. In particular, the deceased’s county of residence, sex, race, age, and level of education are recorded in the files. We use this information to determine both the total number of deaths at the county level as well as the number of deaths among subpopulations of interest. As the mortality files contain no information on the deceased’s income, we use education as a proxy for socioeconomic status.

While the mortality files tell us the number of people who died, they provide us with no information on the size of the population base. When one area reports having more deaths than another, for example, we cannot determine from the mortality files alone whether this is because the population is larger and the death rates are the same, or whether the location experienced a disproportionate number of deaths. To take into account the size of the relevant population, we combine the number of deaths at the county level with linearly interpolated county-month-level population estimates from the 1990, 2000, and 2010 Decennial Censuses.

In addition to population estimates, we also use county-level demographics from both the census and the American Community Survey (ACS) to control for underlying differences across counties and to identify disadvantaged subpopulations. In particular, we consider compositional differences in gender, race, ethnicity, age, education, income, unemployment, rural/urban designation (we define rural counties as those in which over half the population was designated rural from 1990-2010), and poverty across counties when analyzing county-level mortality. As with total population, we linearly interpolate these subpopulation estimates at the county-month-level between the 1990 and 2000 censuses, and either the 2010 census or the 5-year pooled estimates from

the ACS (2008-2012). As shown in Table 2.1, counties in states which allow NPs to prescribe independently at some point during the sample period tend to be less densely populated and less racially diverse. However, both the groups of states have nearly identical employment and poverty statistics.

Self-Reported Mental Health

Our second outcome is the number of days in the past month that a person reports being in poor mental health. This measure comes from the Behavioral Risk Factor Surveillance System survey (BRFSS)—a large, annual phone survey that collects information on health-related risk behaviors, chronic health conditions, and use of preventive services in the US. The BRFSS is representative at the state-year level. Starting in 1993 and in most state-years during our sample frame, respondents were asked the following question:

“Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?”

This question is not designed to draw a particular mental health diagnosis, but rather to indicate whether a respondent experiences any symptoms associated with a wide range of mental health conditions. Importantly, responses are elicited from both those with diagnosed and undiagnosed mental illness, as respondents are not asked whether they have ever been diagnosed with a mental illness by a doctor.

We consider as outcome variables both the number of days reported in poor mental health and an indicator for whether the respondent reported having spent at least 21 of the past 30 days in poor mental health. According to the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5), to be diagnosed with a major depressive episode a patient must have either “a depressed mood most of the day, nearly every day” or “a markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day” for two consecutive weeks. In addition to major depressive disorder, the diagnostic criteria for many mental health conditions include a

specified timeframe over which symptoms must be experienced in order for the diagnosis to apply. Thus, we believe creating a binary variable focusing on those experiencing prolonged symptoms will help identify people suffering from severe forms of mental illness.

On average, BRFSS respondents report spending 3.33 days in the past month in poor mental health, with 67% of respondents reporting no days in poor mental health and 6% of respondents reporting at least 21 days in poor mental health (see Table 2.2). Noticeably, respondents in states which ever had independent prescriptive authority during our sample report fewer days in poor mental health. The BRFSS also includes information on each respondent's sex, race, ethnicity, age, education, income, and employment and health insurance status. These variables allow us both to separately consider disadvantaged populations and to control for underlying differences across respondents in our analysis. As shown in Table 2.2, there are limited differences in the demographics of treatment and control states.

The question on days in poor mental health was not asked in some state-years during our sample frame. These missing state-years correspond to 3.08% of state-year observations and 2.26% of the population. The missing state-years do not correspond with the year before, the year of, or the year after a law change for any state, and thus our identification is not directly affected (see Appendix Table 2.7 for missing state-years).

2.2.3 Health Resources

Increasing the supply of providers who can prescribe medication should impact mental health more for populations living in areas with an insufficient supply of such providers. According to the Health Resources and Services Administration (HRSA), an area is underserved for mental health care services if there is fewer than one psychiatrist for every 30,000 people. Using this definition, we identify underserved counties by combining county-year psychiatrist counts from the HRSA's Area Resource Files with county-year population estimates from the census. As before, we linearly interpolate county-level populations between the 1990, 2000, and 2010 Decennial Censuses.

As the survey data outlined in Section 2.2.2 is at the state level, we also need a measure of how

well-equipped each state is for mental health care services. To take into account the geographic distribution of resources within a state, we define a state as being “underserved” for mental health services in a given year if the population-weighted average of binary, underserved categorizations across all counties in that state is less than the median across all states in that year.

As shown in Figure 2.2, over our sample period, approximately 20% of the US population lived in counties that were underserved for mental health care. Despite yearly fluctuations, this fraction remains fairly stable over our sample period. As shown in Appendix Tables 2.8 and 2.9, underserved counties are on average less densely populated, less educated, and are more white than counties with adequate mental health resources. Refer to Appendix Figures 2.3 and 2.4 for maps of underserved counties and states in 2002, respectively.

2.2.4 Prescription Data

Finally, to examine how prescribing practices of physicians change when NPs gain independent prescriptive authority, we use the Xponent database from IMS Health—a public company specializing in pharmaceutical market intelligence. This dataset contains the number of prescriptions filled for antipsychotics and antidepressants at the physician-month level from 2006 to 2014. Importantly, the data contain both the county and the month of each prescription. These variables allow us to identify whether NPs had independent prescriptive authority when and where the prescription was filled. IMS Health constructs the database by collecting prescription scripts directly from 86% of US retail pharmacies, and estimating the remaining 14% using their patented projection methodology. Unfortunately, we only have prescription data from 2006 onwards. However, eight states granted independent prescriptive authority to NPs during this time frame: Colorado, Hawaii, Maryland, Nevada, North Dakota, Rhode Island, Kentucky, and Minnesota.

2.3 Prescriptive Authority and Mental Health Outcomes

In order to identify whether extending independent prescriptive authority to NPs improves mental health outcomes, we exploit time-series variation in state-level scope of practice legislation and mental health outcomes using a standard difference-in-difference framework. As described in Section 2.2.2, we consider two categories of mental health outcomes: self-reported days in poor mental health and mental-health-related mortality. The impact of prescriptive authority on each category of outcomes is considered in turn below.

2.3.1 Mental-Health-Related Mortality

When NPs are allowed to independently prescribe, do we see reductions in the prevalence of suicides and other mental-health-related deaths? As described in Section 2.2.2, we consider both the number of suicides as well as a broader measure of mental-health-related mortality. Letting $Deaths_{cm}$ denote either of these outcomes in county c in month m , we estimate the following equation:

$$Deaths_{cm} = \beta_0 + \beta_1 Indep. Rx_{sm} + \beta_2 Pop_{cy} + \beta_3 X_{cy} + \gamma_s + \gamma_m + \gamma_y + \epsilon_{cm} \quad (2.1)$$

where $Indep. Rx_{sm}$ is defined a dummy which equals one if NPs had independent prescriptive authority in state s in month m and zero otherwise, $Population_{cy}$ is the population of county c in year y , X_{cy} is a vector of other county-year controls, and λ_s , λ_m , and λ_y are state, month of year, and year fixed effects.¹

To avoid introducing measurement error into the outcome, our preferred specification uses the number of deaths in a county-month as the outcome variable and includes a control for the corresponding population estimate on the right-hand side. While one could use county-level death

¹County-year-level demographics include population, population density, percent male, percent black, percent white, percent hispanic, percent unemployed, percent under 18, percent over 65, percent homeless, percent in poverty, percent with a high school diploma or less, percent with a college degree or more, percent in poverty, median income quintiles, median year structure built, number of people eligible for Medicaid, and the number of practicing psychiatrists and general practitioners.

rates as the outcome variable, death rates are very sensitive to population counts, and precise county-level population estimates are only available every ten years. Whereas measurement error from population estimates on the right-hand side will attenuate our estimated coefficient, it will not affect the precision of our estimated coefficients of interest, as would measurement error in the outcome.

Increasing the supply of providers should impact mental health outcomes most for populations living in areas with an insufficient supply of providers and for populations who find it more difficult to access psychiatrist-provided care. Therefore, we allow the impact of changing scope of practice legislation to differentially influence mental health in counties with an under-provision of mental health care services and for populations who are traditionally disadvantaged. Letting $Underserved_{cy}$ be a dummy which equals one if county c in year y is underserved for mental health care services and zero otherwise, we separately consider counties that are underserved for mental health care services by estimating the following equation:

$$\begin{aligned} Deaths_{cm} = & \beta_0 + \beta_1 Indep. Rx_{sm} + \beta_2 Underserved_{cy} \\ & + \beta_3 Indep. Rx_{sm} \cdot Underserved_{cy} \\ & + \beta_4 Pop_{cy} + \beta_5 X_{cy} + \gamma_s + \gamma_m + \gamma_y + \epsilon_{cm} \end{aligned} \quad (2.2)$$

where all other variables are defined as in Equation 2.1. To look specifically at disadvantaged populations, we again run Equation 2.2 separately for different demographic groups. Results for low education and rural populations are provided in Table 2.3; refer to Appendix Tables 2.10 and 2.11 for results for additional subpopulations.

Table 2.3 demonstrates that independent prescriptive authority for NPs is associated with reductions in both suicides and mental-health-related deaths. Interestingly, we see in Columns 2 - 4 and Columns 6 - 8 that suicides and mental-health-related deaths are actually less prevalent in areas that are underserved by psychiatrists. Despite the lower baselines for these counties, however, they still experience greater percentage reductions than the population as a whole when prescriptive authority is granted to NPs. Comparing Columns 1 and 2, we see that while the average county

does not see any reduction in suicides when prescriptive authority is extended to NPs, counties underserved by psychiatrists experience a reduction of 17%. Columns 5 - 6 paint a similar picture: while on average counties experience no significant reduction in mental-health-related deaths on average when NPs can prescribe, counties that are underserved by psychiatrists experience significant reductions.

In addition, as shown in Columns 3 - 4 and Columns 7 - 8, individuals with low levels of education or residing in rural counties see significant reductions in both the prevalence of suicides and mental-health-related deaths when NPs are granted independent prescriptive authority. For those with low levels of education, suicides and mental-health-related deaths are reduced by 17% and 19%, respectively, when independent prescriptive authority is extended to NPs. The point estimates are smaller but more precisely estimated in rural counties. In particular, suicides and mental-health-related deaths are reduced by 6.4% and 13%, respectively, in rural counties when independent prescriptive authority is extended to NPs. Given that suicides and mental-health-related deaths are quite rare, there is likely more measurement error when we focus on these outcomes in smaller populations than in the population as a whole. As these measurement issues will serve to attenuate our estimates, it is notable that we still identify an effect of comparable magnitude.

Finally, the main results in Table 2.3 are robust to several alternative specifications. In Table 2.3, all regressions are weighted by population size. Appendix Table 2.13 reports the analogous results of unweighted specifications. If anything, the effects of extending prescriptive authority to NPs tend to be more precisely estimated in the unweighted regressions, which makes sense since if the law changes have the biggest impact in less densely populated counties. In addition, the main results are not being driven by any one particular state. Appendix Tables 2.14 and 2.15 show that the point estimates are very stable when separately dropping each state which takes up treatment during our sample period. Finally, Appendix Table 2.16 shows that the results remain when state-level linear time trends are included.

2.3.2 Self-Reported Mental Health

When prescriptive authority is extended to NPs, do we see improvements in the mental health of individuals on a day-to-day basis in conjunction with the decreases in mental-health-related deaths? As described in Section 2.2.2, we consider both the number of days in the past month respondents report being in poor mental health as well as a binary variable which equals one if the respondent reports having spent at least three weeks in poor mental health and zero otherwise. Letting *Poor Mental Health*_{isy} denote either of these outcomes for individual *i* in state *s* in year *y*, we estimate the following equation:

$$Poor\ Mental\ Health_{isy} = \beta_0 + \beta_1 Indep.\ Rx_{sy} + \beta_2 X_{isy} + \gamma_s + \gamma_y + \epsilon_{isy} \quad (2.3)$$

where *Indep. Rx*_{sy} is defined as in Equation 2.1, *X*_{isy} is a vector of individual-level controls, and γ_s and γ_y are state and year fixed effects.² We define a state as having independent prescriptive authority in a given year if NPs had the legislative authority to prescribe medications independently at any point within the year, although all results are robust to alternative timing assumptions,

As before, we consider that extending prescriptive authority to NPs may impact mental health outcomes more for populations living in states with an under-provision of mental health care services and for populations who are traditionally disadvantaged. Letting *Underserved*_{sy} be a dummy which equals one if state *s* in year *y* is less equipped for mental health care services and zero otherwise (as defined in Section 2.2.3), we separately consider states with an under-provision of mental health care services by estimating the following equation:

$$\begin{aligned} Poor\ Mental\ Health_{isy} = & \beta_0 + \beta_1 Indep.\ Rx_{sy} + \beta_2 Underserved_{sy} \\ & + \beta_3 Indep.\ Rx_{sy} \cdot Underserved_{sy} \\ & + \beta_4 X_{isy} + \gamma_s + \gamma_y + \epsilon_{isy} \end{aligned} \quad (2.4)$$

²Individual level controls include sex, age, education, and income quintile dummies, and indicators for race (white, black, and missing), hispanic, employment status, and insurance status.

where all other variables are defined as in Equation 2.3. To look specifically at disadvantaged populations, we estimate the coefficients in Equation 2.4 separately for different subpopulations of interest. Results for individuals with low levels of education are provided in Table 2.4; refer to Appendix Tables 2.17 and 2.18 for results for other subpopulations.

As shown in Table 2.4, independent prescriptive authority for NPs is associated with a significant reduction in the number of poor mental health days reported by survey respondents. Looking first to Column 1, we see that respondents report on average having spent 0.134 fewer days in poor mental health when NPs are allowed to prescribe—a reduction of 4% of the mean. Adhering to the expectation that areas with an insufficient supply of providers before NPs were granted prescriptive authority should experience greater improvements, Column 2 demonstrates that the benefits are concentrated among respondents in states that are less equipped for mental health care. For respondents in underserved states, allowing NPs to prescribe independently is associated with spending 8.4% fewer days in poor mental health—over twice the improvement observed for the population on average.

Furthermore, already disadvantaged populations, see the greatest reductions in poor mental health on a day-to-day basis—here represented by individuals with low levels of education. Column 3 of Table 2.4 shows that individuals who both live in a less equipped state and have low levels of education see even greater improvements. Independent prescriptive authority is associated with a 9.4% reduction in poor mental health days for low education respondents— an additional 12% reduction over the average respondent in a state with an under-provision of psychiatrists, and a 135% reduction over the average respondent in the population as a whole.

In addition, allowing NPs to prescribe medication without physician supervision is associated with a decrease in the likelihood of spending at least three weeks in poor mental health for disadvantaged groups, which is likely indicative of a more severe mental illness. While Columns 4 and 5 of Table 2.4 show that independent prescriptive authority has no impact on this measure of mental health for the population as a whole, individuals who are both living in an underserved state and who have low levels of education experience a 10% reduction in the probability of re-

porting 21+ days in poor mental health when NPs can prescribe. It is possible that in the absence of independent prescriptive authority for NPs, available psychiatrists focus their efforts on those who have more severe mental illnesses. If this were the case, the patterns we observe are to be expected: individuals with minor mental illnesses and individuals with more severe mental illnesses who have difficulty accessing psychiatrist-provided care will benefit the most from the expansion of prescriptive authority. Contrary to what we saw with mental-health-related mortality, the reductions are much larger for disadvantaged groups than those observed across the entire population of individuals living in underserved counties.

Finally, the main results in Table 2.4 are robust to several alternative specifications. Appendix Tables 2.20 and 2.21 show that our results are not being driven by any one state—as before, the point estimates are very stable when separately dropping each state that takes up treatment during our sample period. Unlike the mortality results, however, Appendix Table 2.22 shows that the self-reported mental health results are not robust to adding state-level linear time trends. As all of the variation we are exploiting here is at the state-year level, adding state-level linear time trends in conjunction with state and year fixed effects is asking a lot of the data, so the null result is unsurprising.³

2.4 Prescriptive Authority and Prescribing Patterns

While both the number of NPs seen prescribing and the number of prescriptions written by NPs should increase when they are granted prescriptive authority, the effects of expanded scope of practice on the prescribing behavior of physicians is less clear. Critics note that the prescriptions written by NPs may crowd out the prescriptions previously written by physicians, in which case granting independent prescriptive authority to NPs will have no impact on the overall use of psychotropic treatment. In this case, we should observe a reduction in the number of prescriptions written by

³This is less of a concern when we add state-level linear time trends to the mortality regressions. In Section 2.2.2, the outcome varies at the county-month level. Additionally, there is within-year variation in law changes and within-state variation in health resources.

physicians when prescription authority is extended to NPs. It is also possible, however, that NPs simply meet unmet demand and prescribe for patients who were previously unable to access care. In this case, we should see no impact on the prescribing behavior of physicians when NPs are allowed to independently prescribe. Finally, it could be the case that physicians actually prescribe more medications when NPs are granted independent prescriptive authority. If, for example, NPs bring previously disenfranchised populations into the medical system and refer them to physicians for more specialized care, then we should see the number of prescriptions written by physicians actually increase.

To examine how independent prescriptive authority for NPs influences the prescribing practices of physicians, we exploit within-state variation in both scope of practice laws and the number of prescriptions written by physicians for antipsychotics and antidepressants. We consider the total number of prescriptions written by MDs, the average number of prescriptions written per MD who is observed prescribing, as well as the number of MDs observed prescribing. Letting $Prescriptions_{cm}^{MDs}$ denote one of these outcomes for county c in month m , we estimate the following equation:

$$Prescriptions_{cm}^{MDs} = \beta_0 + \beta_1 Indep. Rx_{sm} + \beta_2 Pop_{cy} + \beta_3 X_{cy} + \gamma_s + \gamma_m + \gamma_y + \epsilon_{cm} \quad (2.5)$$

where $Indep. Rx_{sm}$ is a dummy which equals one if NPs have independent prescriptive authority in state s in month m and zero otherwise, X_{cy} is a vector of county-year controls, and γ_s , γ_m , and γ_y are state, month, and year fixed effects. In order to allow the effect to vary with the provision of local resources, as well as to allow for differential time trends in prescribing habits across areas with a different number of prescribers, we estimate Equation 2.5 separately for counties that are underserved and for counties that are not underserved by psychiatrists.

As shown in Table 2.5, extending prescriptive authority to NPs appears to have no impact on the prescribing behavior of physicians. Neither the total number of prescriptions written for

antidepressants and antipsychotics, the average number of prescriptions per prescribing physician, nor the number of physicians observed prescribing are affected. This holds both in areas that are underserved and in areas that are not underserved by psychiatrists.

2.5 Conclusion

Taken together, our results indicate that granting independent prescriptive authority to NPs is an important policy tool which can be used to improve population mental health. In particular, areas with an under-provision of psychiatrists or with populations who find it difficult to access psychiatrist-provided care can grant independent prescriptive authority to NPs to help mitigate the negative consequences of physician shortages and extend care to disadvantaged populations.

Policies that increase the number of providers who can prescribe medication may be particularly important in the US, where the supply of physicians has not kept pace with rising demand for health care services. Although states often extend prescriptive authority to NPs to address shortages of primary care practitioners with prescribing abilities, we show that these laws also have important implications for mental health. In particular, states that grant independent prescriptive authority to NPs see both improvements in the quality of self-reported mental health as well as reductions in the prevalence of mental-health-related deaths, including suicides. Improvements are greatest for individuals who live in areas that are underserved by psychiatrists and among populations who have been shown to have more difficulty accessing psychiatrist-provided care.

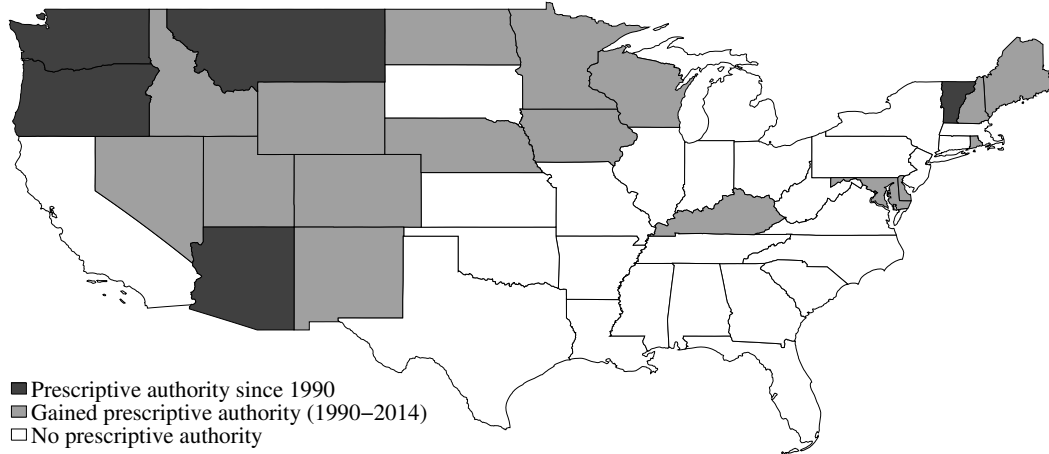
It is noteworthy that we observe a consistent pattern of effects despite using two very different measures of mental health. Self-reported “days in poor mental health” allow us to examine whether the population suffering from mental illnesses of varying severities, including minor mental illnesses, notice improvements when NPs are allowed to prescribe. On the other hand, mental-health-related mortality allows us to examine whether the population suffering from very severe mental illnesses—that is, mental illnesses that may result in death—see improvements when independent prescriptive authority is extended beyond physicians. Even if extending prescriptive

authority to NPs impacts one of these outcomes, it is not clear ex ante that prescriptive authority should also impact the other. In particular, since suicides and other deaths caused by poor mental health are relatively rare, it is possible that population mental health could improve without measurable effects on such extreme outcomes. The consistency of our results across the two categories of outcome measures indicates that prescriptive authority for NPs is associated with improved mental health across a spectrum of severity.

When independent prescriptive authority is extended to NPs, all NPs, not just those who specialize in mental health, have the statutory authority to prescribe. Just like physicians, however, some NPs specialize in psychiatric medicine. Psychiatric NPs with prescriptive authority traditionally provide psychotherapy in addition to psychotropic treatment, unlike the current movement among psychiatrists to only prescribe medications. It is therefore possible that the improvements in mental health that we observe are driven entirely by an increase in “full-service” mental health providers—that is, specialists that provide both psychotherapy and psychotropic treatment. However, it is also possible that our results are driven entirely by an increase in the overall supply of general practitioners who can prescribe. It remains an open question whether extending independent prescriptive authority to NPs results in improved mental health because such laws increase the number of general health care providers who can prescribe psychotropic treatment or because they increase the number of providers who provide psychotherapy in conjunction with psychotropic treatment. Answering this question is a promising area for future research.

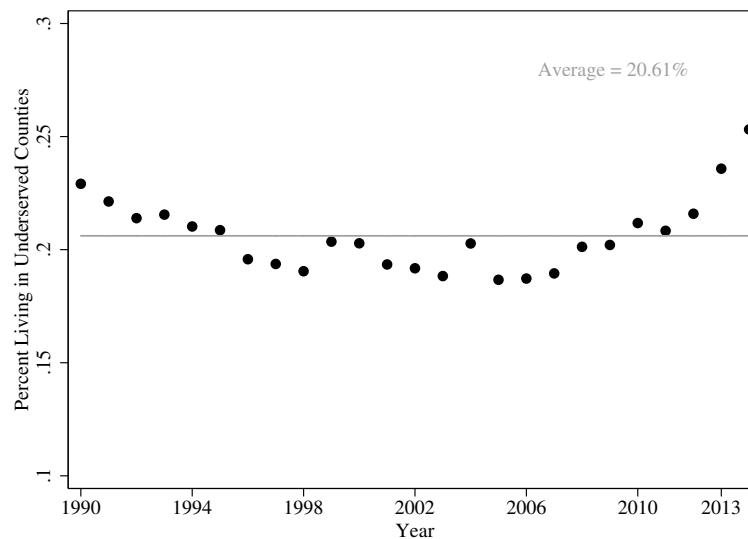
2.6 Figures

Figure 2.1: Changes in Independent Prescriptive Authority for NPs: 1990 – 2014



Notes: NPs in Alaska had independent prescriptive authority before 1990, while NPs in Hawaii gained prescriptive authority in 2010. We define a state as having independent prescriptive authority if NPs registered in the state have the statutory authority to prescribe medications without physician collaboration or supervision.

Figure 2.2: Percent of US Population Living in Underserved Counties: 1990 – 2014



Notes: A county is “underserved” for mental health care services if the county has fewer than one psychiatrist per 30,000 residents. We identify underserved counties by combining county-year information on the number of psychiatrists from the HRSA’s Area Resource Files with county-year population estimates linearly interpolated between the 1990, 2000, and 2010 decennial censuses.

2.7 Tables

Table 2.1: Summary Statistics: County-Level Mortality and Controls, 1990 - 2014

	All	Never Indep. Rx Authority	Ever Indep. Rx Authority
Deaths per 100,000:			
Suicides	0.99	0.93	1.17
All Mental-Health-Related Causes	1.85	1.77	2.11
Independent Prescriptive Authority	0.13	0	0.61
Low Psychiatrist-to-Pop. Ratio	0.21	0.20	0.21
Total Population	1,039,476	1,173,171	550,514
Population Density (Per Sq. Mile)	2,209	2,247	729
Percent Male	0.49	0.49	0.49
Percent White	0.75	0.73	0.81
Percent Black	0.12	0.14	0.06
Percent Hispanic	0.13	0.14	0.10
Percent Unemployed	0.06	0.06	0.06
Percent Poverty	0.25	0.25	0.25
Percent 18 and Under	0.13	0.13	0.12
Percent 65 and Over	0.47	0.48	0.43
Percent HS or Less	0.28	0.29	0.27
Median Income: 1st Quintile	0.05	0.06	0.03
Median Income: 2nd Quintile	0.08	0.08	0.06
Median Income: 3rd Quintile	0.12	0.13	0.11
Median Income: 4th Quintile	0.21	0.21	0.24
Median Income: 5th Quintile	0.54	0.53	0.56
Percent College or More	0.25	0.25	0.26
Median Year Structure Built	1970	1969	1973
Medicaid Eligible per 100,000	18,724	19,347	16,447
Psychiatrists per 100,000	12	12	11
Primary Care MDs per 100,000	77	77	78
Rural County	0.13	0.13	0.15
Observations	925,174	640,882	284,292

Notes: Data is at the county-year level, weighted by population. “Mental-Health-Related Deaths” includes suicide, deaths of unknown intent, and accidental deaths involving firearms, trains, and poisonings. “Ever (Never) Independent Rx Authority” includes respondents living in states which had independent prescriptive authority for NPs at some point (at no point) during our sample. “Independent Prescriptive Authority” and “Low Psychiatrist-to-Pop. Ratio” reflect whether counties allowed NPs independent prescriptive authority or were underserved for mental health services in a given year. Mortality statistics are taken from the US Mortality Files, provider counts from the the HRSA’s Area Resource Files, and all other variables from the 1990, 2000, and 2010 decennial censuses and the 5-year pooled (2008-2012) American Community Survey (ACS). Census and ACS variables are linearly interpolated at the county-year level.

Table 2.2: Summary Statistics: Self-Reported Mental Health and Controls, 1993 - 2013

	All	Never Indep. Rx Authority	Ever Indep. Rx Authority
Days in Poor Mental Health:			
Average	3.33	3.36	3.16
Percent 0	0.67	0.67	0.66
Percent >= 21	0.06	0.06	0.05
Independent Prescriptive Authority	0.14	0	0.75
Low Psychiatrist-to-Population Ratio	0.37	0.39	0.26
Male	0.48	0.48	0.49
White	0.71	0.70	0.78
Black	0.10	0.11	0.05
Hispanic	0.12	0.13	0.09
Employed	0.52	0.52	0.54
Health Insurance	0.84	0.84	0.86
Age: 18 to 34	0.31	0.31	0.32
Age: 35 to 44	0.20	0.20	0.20
Age: 45 to 54	0.18	0.18	0.18
Age: 55 to 64	0.13	0.13	0.13
Age: 65 and Over	0.17	0.17	0.17
Education: HS or Less	0.43	0.44	0.39
Education: Some College or More	0.57	0.56	0.60
Income: 1st Quintile	0.20	0.20	0.17
Income: 2nd Quintile	0.17	0.17	0.18
Income: 3rd Quintile	0.17	0.17	0.18
Income: 4th Quintile	0.18	0.18	0.19
Income: 5th Quintile	0.14	0.14	0.15
Observations	5,670,468	3,399,048	2,271,420

Notes: Data is at the individual level, weighted using BRFSS sample weights. “Days in Poor Mental Health” was not asked in WY in 1993, RI in 1994, DC in 1995, 29 states in 2002, and HI in 2004. “Ever (Never) Independent Rx Authority” includes respondents living in states which had independent prescriptive authority for NPs at some point (at no point) during our sample. “Independent Prescriptive Authority” and “Low Psychiatrist-to-Pop. Ratio” reflect the fraction of respondents living in a state with independent prescriptive authority or that was underserved for mental health services in the year that he/she responded. Some categorical variables do not sum to one; the difference reflects the percentage of missings.

Table 2.3: Independent Prescriptive Authority and Mental-Health-Related Mortality

	Suicides				Mental-Health-Related Deaths			
	(1) Full Sample	(2) Full Sample	(3) Low Educ.	(4) Rural Counties	(5) Full Sample	(6) Full Sample	(7) Low Educ.	(8) Rural Counties
Indep. Prescriptive Authority	0.143 (0.338)	0.624 (0.393)	0.444 (0.362)	0.033 (0.022)	-0.825 (0.564)	-0.037 (0.592)	0.209 (0.515)	0.028 (0.051)
Low Psychiatrist-to-Pop. Ratio		-0.598** (0.266)	-0.519** (0.218)	-0.020** (0.008)		-0.843* (0.473)	-0.826** (0.409)	-0.031 (0.020)
Interaction		-2.209* (1.284)	-1.329* (0.740)	-0.065** (0.027)		-3.611* (2.102)	-2.131* (1.206)	-0.130** (0.050)
Observations	925,174	925,174	925,174	526,294	925,174	925,174	925,174	526,294
R^2	0.920	0.921	0.863	0.205	0.936	0.936	0.902	0.313
Mean Dependent Variable	9.51	9.51	5.23	0.50	17.12	17.12	10.23	0.80

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. "Low Educ" is defined as having a high school degree or less. Columns 3 and 6 further control for the size of this low education population. Refer to Table 2.12 for full regression results.

Table 2.4: Independent Prescriptive Authority and Self-Reported Mental Health

	Days in Poor Mental Health			21+ Days in Poor Mental Health		
	(1) Full sample	(2) Full sample	(3) Low Educ.	(4) Full sample	(5) Full sample	(6) Low Educ.
Indep. Prescriptive Authority	-0.134* (0.072)	-0.093 (0.083)	-0.095 (0.136)	-0.003 (0.002)	-0.002 (0.002)	-0.002 (0.004)
Low Psychiatrist-to-Pop. Ratio		0.113 (0.075)	0.157 (0.124)		0.002** (0.001)	0.004* (0.002)
Interaction		-0.186* (0.109)	-0.267* (0.140)		-0.003 (0.002)	-0.005** (0.002)
Observations	5,670,468	5,670,468	2,296,282	5,670,468	5,670,468	2,296,282
R^2	0.044	0.044	0.044	0.024	0.024	0.025
Mean Dependent Variable	3.33	3.33	3.86	0.06	0.06	0.07

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. Additional controls include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles are also included. “Low Educ” is defined as having a high school degree, less than a high school degree, or missing education information. Refer to Table 2.19 for the full regression results.

Table 2.5: Independent Prescriptive Authority and Prescriptions by MDs

	Not Underserved Counties			Underserved Counties		
	(1) Total	(2) Average	(3) N Prescribers	(4) Total	(5) Average	(6) N Prescribers
Indep. Prescriptive Authority	-2.648 (2.996)	-0.157 (2.101)	1.297 (8.875)	-0.504 (0.454)	-4.072 (2.974)	-1.764 (1.239)
Observations	114,288	114,288	114,288	224,184	224,184	224,184
R^2	0.910	0.299	0.933	0.843	0.199	0.882
Mean Dependent Variable	143.52	65.31	462.35	13.63	63.43	33.94

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. “Total” refers to the total number of prescriptions written by MDs for antidepressants and antipsychotics in hundreds. “Average” refers to the average number of scripts written for these medications across doctors observed prescribing drugs in these classes. “N Prescribers” refers to the number of unique prescribers observed prescribing drugs in these classes.

2.8 Appendix

Data Appendix

Compiling the prescriptive authority database

The Nurse Practitioner is a journal addressing clinical issues relevant to NPs and other primary care providers. Every January since 1989, the journal has published the “Annual Legislative Update” which summarizes both the practice environment and the level of prescriptive authority for NPs in each state. While informative, these overviews do not consistently include dates of legislative action nor comprehensive coverage of the precise changes made to a state’s legislation. Therefore, the information provided by the journal alone is not sufficient for a quantitative analysis of independent prescriptive authority.

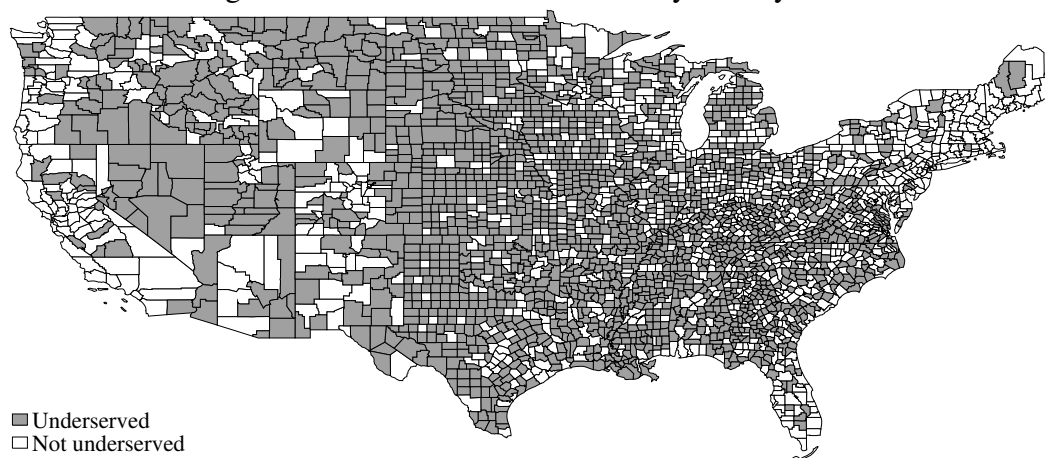
In many states, NPs have the statutory authority to prescribe only under the direct supervision of a physician or with a collaborative practice agreement. While such laws should also increase access to medication over a regime in which NPs have no ability to prescribe, we expect their impacts to be more limited than granting NPs full independent prescriptive authority. Anecdotal evidence suggests that NPs have difficulty finding physicians who are willing to supervise or work in collaboration, especially in areas with few physicians. Therefore, collaborate practice agreements and mandatory supervision likely limit the expansion of access to medication that could be achieved by allowing NPs to prescribe independently, particularly in places where more prescribers may be needed the most.

Missing state-years in the BRFSS data

The BRFSS did not ask the mental health question in Wyoming in 1993, Rhode Island in 1994, D.C. in 1995, 29 states in 2002, and Hawaii in 2004 (see Table 2.7).

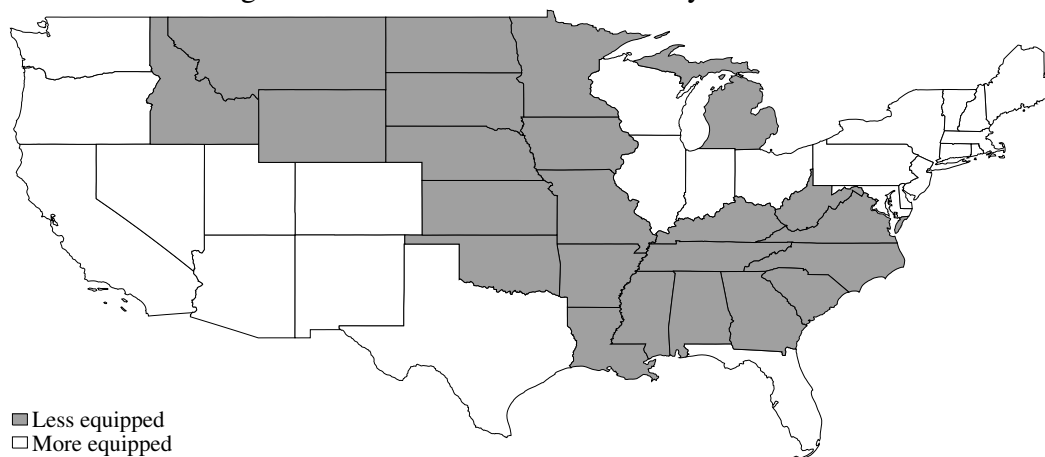
2.8.1 Appendix Figures

Figure 2.3: Mental Health Services by County: 2002



Notes: Following the definition provided by the HRSA, a county is “underserved” for mental health care services in a given year if the county has fewer than one psychiatrist per 30,000 residents. We identify underserved counties by combining county-year information on the number of psychiatrists from the HRSA’s Area Resource Files with county-year population estimates linearly interpolated between the 1990, 2000, and 2010 decennial censuses.

Figure 2.4: Mental Health Services by State: 2002



Notes: A state is “less equipped” for mental health services in a given year if the population-weighted average of binary, underserved categorizations across all counties in that state is less than the median across all states in that year. To construct the binary, underserved categorizations at the county-level, we use the definition provided by the HRSA: a county is “underserved” for mental health care services if the county has fewer than one psychiatrist per 30,000 residents. We identify underserved counties by combining county-year information on the number of psychiatrists from the HRSA’s Area Resource Files with county-year population estimates linearly interpolated between the 1990, 2000, and 2010 decennial censuses.

2.8.2 Appendix Tables

Table 2.6: Robustness: Correlates of Law Changes

	Unemployment Rates		No. of Psychiatrists		No. of General Practitioners	
	(1)	(2)	(3)	(4)	(5)	(6)
	FE & Pop. Only All Controls		FE & Pop. Only All Controls		FE & Pop. Only All Controls	
Indep. Prescriptive Authority	-0.000	0.000	8.109	8.185*	-52.965*	-31.032
	(0.002)	(0.002)	(7.339)	(4.549)	(28.158)	(19.434)
Low Psychiatrist-to-Pop. Ratio	-0.000	0.000	3.786	12.926**	-83.379**	26.667
	(0.000)	(0.000)	(5.801)	(5.780)	(35.827)	(17.914)
Interaction	0.000	-0.000	4.930	10.498	53.362	-5.275
	(0.000)	(0.001)	(14.896)	(9.822)	(73.743)	(49.572)
Observations	925,174	925,174	925,174	925,174	925,174	925,174
R^2	0.850	0.851	0.890	0.944	0.967	0.986

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state and year fixed effects. The full controls specifications also include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles (minus the dependent variable). “Low Educ” is defined as having a high school degree or less.

	General Health		Insured	
	(1)	(2)	(3)	(4)
	FE Only	All Controls	FE Only	All Controls
Indep. Prescriptive Authority	-0.006	-0.001	0.007	0.003
	(0.007)	(0.007)	(0.005)	(0.005)
Low Psychiatrist-to-Pop. Ratio	0.091*	0.106***	0.012	0.003
	(0.046)	(0.036)	(0.035)	(0.032)
Interaction	-0.014	0.000	-0.002	-0.005
	(0.009)	(0.007)	(0.004)	(0.004)
Observations	6,099,200	6,099,200	6,118,736	6,118,736
R^2	0.009	0.165	0.015	0.172

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. The full controls specifications also include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles (minus the dependent variable). “Low Educ” is defined as having a high school degree, less than a high school degree, or missing education information. “Low Inc” is defined as being in the bottom two quintiles of income.

Table 2.7: Missing BRFSS State-Years

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014 mi
AK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AR	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AZ	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CO	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DC	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ID	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IN	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KS	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KY	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ME	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MN	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MO	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MS	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ND	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NH	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NJ	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NV	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NY	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OH	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OR	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SC	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TN	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TX	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
UT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WV	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
WY	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Total Missings	1	1	1	0	0	0	0	0	0	29	0	1	0	0	0	0	0	0	0	0	0	0

Table 2.8: Summary Statistics: Self-Reported Mental Health and Controls, 1993 - 2013

	Never Indep. Prescript. Authority				Ever Indep. Prescript. Authority		
	All	All	Never under-served	Ever under-served	All	Never under-served	Ever under-served
Days in Poor Mental Health:							
Average	3.33	3.36	3.39	3.34	3.16	3.16	3.17
Percent 0	0.67	0.67	0.65	0.68	0.66	0.67	0.65
Percent >= 21	0.06	0.06	0.06	0.06	0.05	0.05	0.05
Independent Prescriptive Authority	0.14	0	0	0	0.75	0.53	0.94
Low Psychiatrist-to-Population Ratio	0.37	0.39	0	0.69	0.26	0	0.47
Male	0.48	0.48	0.48	0.48	0.49	0.49	0.49
White	0.71	0.70	0.64	0.74	0.78	0.70	0.85
Black	0.10	0.11	0.08	0.13	0.05	0.09	0.02
Hispanic	0.12	0.13	0.19	0.08	0.09	0.12	0.07
Employed	0.52	0.52	0.51	0.53	0.54	0.54	0.53
Health Insurance	0.84	0.84	0.85	0.84	0.86	0.86	0.86
Age: 18 to 34	0.31	0.31	0.31	0.32	0.32	0.31	0.32
Age: 35 to 44	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Age: 45 to 54	0.18	0.18	0.18	0.18	0.18	0.18	0.18
Age: 55 to 64	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Age: 65 and Over	0.17	0.17	0.18	0.17	0.17	0.16	0.17
Education: HS or Less	0.43	0.44	0.42	0.45	0.39	0.39	0.40
Education: Some College or More	0.57	0.56	0.58	0.54	0.60	0.61	0.60
Income: 1st Quintile	0.20	0.20	0.21	0.20	0.17	0.16	0.18
Income: 2nd Quintile	0.17	0.17	0.16	0.18	0.18	0.17	0.20
Income: 3rd Quintile	0.17	0.17	0.16	0.17	0.18	0.17	0.19
Income: 4th Quintile	0.18	0.18	0.19	0.18	0.19	0.19	0.19
Income: 5th Quintile	0.14	0.14	0.15	0.13	0.15	0.18	0.13
Observations	5,670,468	3,399,048	1,104,585	2,294,463	2,271,420	973,174	1,298,246

Notes: Raw data is at the individual level, weighted using sample weights. “Days in Poor Mental Health” reflects responses to the question “Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?” This question was asked on the BRFSS in all states in 1997 – 2001, 2003 and 2005 – 2013. It was not asked in Wyoming in 1993, Rhode Island in 1994, D.C. in 1995, 29 states in 2002, and Hawaii in 2004. “Ever Independent Prescriptive Authority” reflects whether the respondent was living in a state which had independent prescriptive authority for NPs at some point during our sample. Similarly, “Ever Underserved” reflects whether the respondent was living in a state which was less equipped for mental health care services at some point during our sample. “Independent Prescriptive Authority” reflects whether the respondent was living in a state with independent prescriptive authority in the year that he/she responded. Similarly, “Low Psychiatrist-to-Pop. Ratio” reflects whether the respondent was living in a state that was less equipped for mental health services in the year that he/she responded. Construction of both “Independent Prescriptive Authority” and “Low Psychiatrist-to-Pop. Ratio” requires information not provided in the BRFSS; their construction is outlined in Sections 2.2.1 and 2.2.3, respectively. All other variables come from the BRFSS and were asked in conjunction with the question on days in poor mental health. Some categorical variables do not sum to one; the difference reflects the percentage of missings.

Table 2.9: Summary Statistics: County-Level Mortality and Controls, 1990 - 2014

	Never Indep. Prescrip. Authority				Ever Indep. Prescript. Authority		
	All	All	Never under-served	Ever under-served	All	Never under-served	Ever under-served
Deaths per 100,000:							
Suicides	0.99	0.93	0.87	1.07	1.17	1.14	1.22
All Mental-Health-Related Causes	1.85	1.77	1.69	1.93	2.11	2.10	2.14
Independent Prescriptive Authority	0.13	0	0	0	0.61	0.62	0.61
Low Psychiatrist-to-Pop. Ratio	0.21	0.20	0	0.59	0.21	0	0.51
Total Population	1,039,476	1,173,171	1,699,192	169,328	550,514	838,276	142,774
Population Density (Per Sq. Mile)	2,209	2,247	3,359	332	729	982	369
Percent Male	0.49	0.49	0.49	0.49	0.49	0.49	0.50
Percent White	0.75	0.73	0.69	0.82	0.81	0.78	0.86
Percent Black	0.12	0.14	0.15	0.11	0.06	0.08	0.04
Percent Hispanic	0.13	0.14	0.17	0.08	0.10	0.11	0.08
Percent Unemployed	0.06	0.06	0.06	0.06	0.06	0.06	0.05
Percent Poverty	0.25	0.25	0.25	0.25	0.25	0.24	0.26
Percent 18 and Under	0.13	0.13	0.12	0.14	0.12	0.12	0.13
Percent 65 and Over	0.47	0.48	0.44	0.55	0.43	0.40	0.48
Percent HS or Less	0.28	0.29	0.29	0.29	0.27	0.27	0.27
Median Income: 1st Quintile	0.05	0.06	0.01	0.14	0.03	0	0.07
Median Income: 2nd Quintile	0.08	0.08	0.04	0.16	0.06	0.02	0.13
Median Income: 3rd Quintile	0.12	0.13	0.10	0.19	0.11	0.06	0.17
Median Income: 4th Quintile	0.21	0.21	0.20	0.22	0.24	0.25	0.22
Median Income: 5th Quintile	0.54	0.53	0.65	0.30	0.56	0.66	0.42
Percent College or More	0.25	0.25	0.29	0.18	0.26	0.30	0.21
Median Year Structure Built	1970	1969	1967	1973	1973	1973	1973
Medicaid Eligible per 100,000	18,724	19,347	19,942	18,210	16,447	15,821	17,334
Psychiatrists per 100,000	12	12	17	4	11	15	5
Primary Care MDs per 100,000	77	77	92	48	78	93	56
Rural County	0.13	0.13	0.01	0.36	0.15	0.04	0.31
Observations	925,174	640,882	105,300	535,582	284,292	44,984	239,308

Notes: Data is at the county-year level, weighted by population. “Suicides” and “Mental-Health-Related Deaths” are taken from the US Mortality Files. “Mental-Health-Related Deaths” includes suicides, deaths of unknown intent, and accidental deaths (accidental deaths involving firearms, trains, and poisonings). “Ever Independent Prescriptive Authority” reflects whether NPs ever had independent prescriptive authority in the state. Similarly, “Ever Underserved” reflects whether the county was ever underserved for mental health services. “Independent Prescriptive Authority” reflects whether NPs had independent prescriptive authority in a given county-year. The construction of this variable is described in Section 2.2.1. Similarly, “Low Psychiatrist-to-Pop. Ratio” reflects whether the county was underserved for mental health services in that year. The construction of this variable is described in Section 2.2.3. “Number of Practicing Psychiatrists” is provided by the AMA to the HRSA and is accessed by the authors from the HRSA’s Area Resource Files. All other variables are from the 1990, 2000, and 2010 decennial censuses and the 5-year polled (2008-2012) American Community Survey (ACS). For census and ACS variables, years other than 1990, 2000, and 2010 are linearly interpolated at the county-year level.

Table 2.10: Prescriptive Authority and Suicides

	(1) Full sample	(2) Black	(3) Middle Age	(4) Under 18	(5) Low Educ.	(6) Male	(7) Female	(8) Rural
Indep. Prescriptive Authority	0.624 (0.393)	-0.069 (0.152)	0.060 (0.093)	0.036 (0.024)	0.444 (0.362)	0.516* (0.303)	0.160* (0.094)	0.033 (0.022)
Low Psychiatrist-to-Pop. Ratio	-0.598** (0.266)	-0.046 (0.045)	-0.104** (0.045)	-0.026 (0.016)	-0.519** (0.218)	-0.471** (0.205)	-0.102* (0.054)	-0.020** (0.008)
Interaction	-2.209* (1.284)	-0.085* (0.045)	-0.696** (0.340)	-0.067 (0.049)	-1.329* (0.740)	-1.521 (0.914)	-0.559* (0.317)	-0.065** (0.027)
Observations	925,174	925,174	925,174	925,174	925,174	925,174	925,174	526,294
R^2	0.921	0.686	0.824	0.422	0.863	0.914	0.796	0.205
Mean Dependent Variable	9.51	0.95	2.53	0.32	5.23	7.29	2.22	0.50

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. Each column further controls for the size of the population for which suicides are being measured (e.g. Column 2 additionally controls for the size of the black population). “Low Educ” is defined as having a high school degree or less.

Table 2.11: Prescriptive Authority and Mental-Health-Related Mortality

	(1) Full sample	(2) Black	(3) Middle Age	(4) Under 18	(5) Low Educ.	(6) Male	(7) Female	(8) Rural
Indep. Prescriptive Authority	-0.037 (0.592)	0.044 (0.168)	-0.029 (0.213)	0.116* (0.061)	0.209 (0.515)	0.164 (0.438)	-0.136 (0.178)	0.028 (0.051)
Low Psychiatrist-to-Pop. Ratio	-0.843* (0.473)	-0.050 (0.078)	-0.173* (0.102)	-0.081* (0.048)	-0.826** (0.409)	-0.625* (0.355)	-0.185* (0.106)	-0.031 (0.020)
Interaction	-3.611* (2.102)	-0.394** (0.174)	-1.230* (0.618)	-0.090 (0.127)	-2.131* (1.206)	-2.448 (1.513)	-0.999* (0.527)	-0.130** (0.050)
Observations	925,174	925,174	925,174	925,174	925,174	925,174	925,174	526,294
R^2	0.936	0.858	0.885	0.581	0.902	0.932	0.865	0.313
Mean Dependent Variable	17.12	2.41	4.96	0.69	10.23	12.78	4.34	0.80

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. Each column further controls for the size of the population for which mental-health-related mortality is being measured (e.g. Column 2 additionally controls for the size of the black population). “Low Educ” is defined as having a high school degree or less.

Table 2.12: Mental-Health-Related Mortality: Full Regression Results

	Suicides				Mental-Health-Related Deaths			
	(1) Full Sample	(2) Full Sample	(3) Low Educ.	(4) Rural Counties	(5) Full Sample	(6) Full Sample	(7) Low Educ.	(8) Rural Counties
Indep. Prescribe	0.143 (0.338)	0.624 (0.393)	0.444 (0.362)	0.033 (0.022)	-0.825 (0.564)	-0.037 (0.592)	0.209 (0.515)	0.028 (0.051)
Low Psych./Pop.		-0.598** (0.266)	-0.519** (0.218)	-0.020** (0.008)		-0.843* (0.473)	-0.826** (0.409)	-0.031 (0.020)
Interaction		-2.209* (1.284)	-1.329* (0.740)	-0.065** (0.027)		-3.611* (2.102)	-2.131* (1.206)	-0.130** (0.050)
Total Population	0.000*** (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000*** (0.000)
Pop. Density	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.001* (0.000)
Percent Male	-22.45*** (8.089)	-16.430** (8.152)	-8.143 (6.111)	-0.510** (0.232)	-38.41*** (14.079)	-29.118** (14.119)	-15.121 (11.809)	-0.683 (0.483)
Percent White	4.248 (3.995)	3.987 (3.560)	1.834 (2.371)	-0.019 (0.119)	10.850 (6.551)	10.373* (5.861)	4.836 (3.601)	0.150 (0.157)
Percent Black	11.659*** (4.285)	11.017*** (3.989)	7.198** (3.052)	-0.258** (0.109)	22.499*** (7.229)	21.485*** (6.754)	14.508*** (5.150)	-0.378** (0.153)
Percent Hispanic	0.198 (1.519)	-0.287 (1.430)	-0.837 (1.001)	-0.218*** (0.066)	0.963 (2.719)	0.246 (2.571)	-0.653 (1.606)	-0.314** (0.125)
Percent Unempl.	35.226** (13.196)	35.458** (13.338)	16.120** (6.025)	-0.278 (0.437)	80.868** (30.234)	81.188** (30.540)	43.575** (16.817)	-0.531 (1.176)
Percent Under 18	-18.55*** (6.861)	-15.081** (7.025)	-2.697 (3.990)	-0.687*** (0.224)	-22.702* (12.179)	-17.303 (12.663)	0.303 (6.763)	-1.344** (0.566)
Percent Over 65	-8.024 (4.850)	-5.866 (4.033)	-2.934 (3.470)	0.343 (0.216)	-22.131** (10.386)	-18.660** (9.206)	-9.945 (7.747)	0.350 (0.471)
Percent HS or Less	-0.910 (4.967)	0.070 (4.544)	-3.407 (3.285)	-0.231 (0.142)	4.372 (7.735)	5.673 (7.291)	-4.984 (4.495)	-0.562** (0.264)
Percent Poverty	-5.059** (1.978)	-5.021** (2.105)	-0.833 (2.200)	-0.014 (0.084)	-8.084** (3.518)	-7.986** (3.704)	0.320 (3.446)	0.216 (0.210)
Median Inc.: Q2	0.490 (0.317)	0.484* (0.288)	0.261 (0.249)	0.009 (0.011)	1.068*** (0.380)	1.070*** (0.319)	0.542* (0.281)	-0.012 (0.023)
Median Inc.: Q3	0.409 (0.312)	0.341 (0.269)	0.033 (0.238)	0.014 (0.013)	0.764 (0.647)	0.675 (0.575)	0.135 (0.503)	-0.004 (0.027)
Median Inc.: Q4	1.483*** (0.326)	1.371*** (0.272)	0.599** (0.269)	0.007 (0.018)	2.264** (0.981)	2.106** (0.898)	0.843 (0.797)	-0.020 (0.039)
Median Inc.: Q5	1.380** (0.538)	1.284** (0.482)	0.714 (0.445)	-0.018 (0.023)	1.410 (1.286)	1.272 (1.208)	0.602 (1.063)	-0.058 (0.049)
Percent College +	-0.201 (5.893)	-0.424 (5.558)	-4.299 (3.458)	-0.759*** (0.234)	2.038 (10.176)	1.619 (9.679)	-9.699** (4.483)	-1.668*** (0.433)
Median Year Built	-0.015 (0.033)	-0.014 (0.032)	-0.008 (0.027)	0.003*** (0.001)	-0.035 (0.048)	-0.034 (0.047)	-0.018 (0.041)	0.004*** (0.001)
Medicaid Eligible	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)
Psychiatrists	0.007*** (0.002)	0.007*** (0.002)	0.005*** (0.001)	-0.006*** (0.002)	0.000 (0.006)	0.001 (0.006)	0.001 (0.003)	-0.011*** (0.004)
Primary Care MDs	-0.004*** (0.001)	-0.004*** (0.001)	-0.002* (0.001)	0.001** (0.001)	-0.001 (0.003)	-0.001 (0.003)	0.001 (0.002)	0.003*** (0.001)
Observations	925,174	925,174	925,174	526,294	925,174	925,174	925,174	526,294
R ²	0.920	0.921	0.863	0.205	0.936	0.936	0.902	0.313
Mean Dep. Var.	9.51	9.51	5.23	0.50	17.12	17.12	10.23	0.80

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. “Low Educ” is defined as having a high school degree or less. Columns 3 and 6 further control for the size of this low education population.

Table 2.13: Prescriptive Authority and Mortality: Unweighted

	Suicides				Mental-Health-Related Deaths			
	(1) Full Sample	(2) Full Sample	(3) Low Educ.	(4) Rural Counties	(5) Full Sample	(6) Full Sample	(7) Low Educ.	(8) Rural Counties
Indep. Prescribe	-0.021 (0.047)	0.252** (0.125)	0.125 (0.078)	0.020 (0.017)	-0.203* (0.110)	0.204 (0.201)	0.065 (0.123)	0.006 (0.031)
Low Psych.-to-Pop. Ratio		-0.127*** (0.040)	-0.095*** (0.030)	-0.013** (0.007)		-0.152** (0.069)	-0.129** (0.058)	-0.023* (0.012)
Interaction		-0.402** (0.173)	-0.244** (0.099)	-0.043** (0.018)		-0.599** (0.261)	-0.351** (0.146)	-0.068** (0.031)
Observations	925,174	925,174	925,174	526,294	925,174	925,174	925,174	526,294
R^2	0.808	0.809	0.693	0.178	0.854	0.855	0.787	0.273
Mean Dependent Variable	1.02	1.02	0.61	0.28	1.70	1.70	1.08	0.45

Notes: Observations are at the county-month level, and are not population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. “Low Educ” is defined as having a high school degree or less. Columns 3 and 6 further control for the size of this low education population.

Table 2.14: Suicides Robustness (Leave Out)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ID	NV	UT	CO	WY	NE	NM	ND	MI
Indep. Prescriptive Authority	0.568 (0.390)	0.685* (0.407)	0.487 (0.356)	0.648 (0.468)	0.634 (0.396)	0.656 (0.410)	0.571 (0.388)	0.625 (0.396)	0.664* (0.393)
Low Psychiatrist-to-Pop. Ratio	-0.602** (0.268)	-0.549** (0.261)	-0.593** (0.264)	-0.602** (0.270)	-0.598** (0.267)	-0.598** (0.267)	-0.604** (0.266)	-0.604** (0.267)	-0.612** (0.275)
Interaction	-2.301* (1.324)	-2.257* (1.294)	-1.948 (1.365)	-2.234* (1.315)	-2.266* (1.306)	-2.345* (1.349)	-2.272* (1.353)	-2.206* (1.287)	-2.221* (1.293)
Observations	912,505	920,399	916,948	907,607	918,317	899,838	915,511	910,280	900,308
R^2	0.921	0.922	0.922	0.921	0.921	0.921	0.921	0.921	0.922
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	IA	WI	KY	MD	DE	RI	CT	ME	HI
Indep. Prescriptive Authority	0.683 (0.411)	0.683 (0.437)	0.624 (0.400)	0.945** (0.373)	0.633 (0.400)	0.652 (0.402)	0.626 (0.395)	0.610 (0.398)	0.604 (0.404)
Low Psychiatrist-to-Pop. Ratio	-0.602** (0.264)	-0.615** (0.268)	-0.598** (0.271)	-0.505** (0.246)	-0.599** (0.267)	-0.595** (0.266)	-0.597** (0.268)	-0.598** (0.266)	-0.601** (0.267)
Interaction	-2.608* (1.384)	-2.592* (1.432)	-2.217* (1.295)	-2.272* (1.311)	-2.210* (1.285)	-2.210* (1.284)	-2.207* (1.284)	-2.252* (1.302)	-2.205* (1.287)
Observations	895,478	903,589	889,211	917,974	924,274	923,674	922,774	920,374	923,943
R^2	0.921	0.921	0.921	0.924	0.921	0.921	0.921	0.921	0.921

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. Each column further controls for the size of the population for which suicide mortality is being measured (e.g. Column 2 additionally controls for the size of the black population). "Low Educ" is defined as having a high school degree or less.

Table 2.15: Mental-Health-Related Mortality Robustness (Leave Out)

	(1) ID	(2) NV	(3) UT	(4) CO	(5) WY	(6) NE	(7) NM	(8) ND	(9) MI
Indep. Prescriptive Authority	-0.032 (0.618)	-0.068 (0.600)	-0.137 (0.589)	-0.010 (0.715)	-0.025 (0.597)	0.079 (0.625)	-0.181 (0.575)	-0.030 (0.600)	0.000 (0.598)
Low Psychiatrist-to-Pop. Ratio	-0.851* (0.477)	-0.779 (0.468)	-0.841* (0.472)	-0.858* (0.479)	-0.844* (0.474)	-0.849* (0.474)	-0.868* (0.470)	-0.854* (0.475)	-0.957** (0.471)
Interaction	-3.725* (2.169)	-3.633* (2.121)	-3.474 (2.262)	-3.638* (2.152)	-3.694* (2.137)	-3.876* (2.202)	-3.593 (2.221)	-3.602* (2.107)	-3.481 (2.107)
Observations	912,505	920,399	916,948	907,607	918,317	899,838	915,511	910,280	900,308
R^2	0.936	0.937	0.936	0.936	0.936	0.936	0.936	0.936	0.937
	(1) IA	(2) WI	(3) KY	(4) MD	(5) DE	(6) RI	(7) CT	(8) ME	(9) HI
Indep. Prescriptive Authority	0.062 (0.619)	-0.063 (0.676)	-0.074 (0.597)	0.542 (0.470)	-0.033 (0.606)	-0.051 (0.598)	-0.034 (0.596)	-0.067 (0.611)	-0.094 (0.609)
Low Psychiatrist-to-Pop. Ratio	-0.851* (0.470)	-0.883* (0.478)	-0.843* (0.480)	-0.707 (0.452)	-0.844* (0.475)	-0.843* (0.473)	-0.837* (0.477)	-0.846* (0.473)	-0.847* (0.474)
Interaction	-4.224* (2.271)	-4.115* (2.382)	-3.611* (2.121)	-3.810* (2.145)	-3.612* (2.105)	-3.606* (2.104)	-3.610* (2.101)	-3.662* (2.132)	-3.596* (2.107)
Observations	895,478	903,589	889,211	917,974	924,274	923,674	922,774	920,374	923,943
R^2	0.936	0.936	0.936	0.937	0.936	0.936	0.936	0.936	0.936

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. Each column further controls for the size of the population for which mental-health-related mortality is being measured (e.g. Column 2 additionally controls for the size of the black population). “Low Educ” is defined as having a high school degree or less.

Table 2.16: Mental-Health-Related Mortality: Including State-Specific Linear Time Trends

	Suicides				Mental-Health-Related Deaths			
	(1) Full Sample	(2) Full Sample	(3) Low Educ.	(4) Rural Counties	(5) Full Sample	(6) Full Sample	(7) Low Educ.	(8) Rural Counties
Indep. Prescriptive Authority	-0.478 (0.527)	-0.029 (0.598)	-0.285 (0.427)	-0.003 (0.036)	-0.666 (0.584)	0.074 (0.736)	-0.378 (0.459)	0.029 (0.047)
Low Psychiatrist-to-Pop. Ratio		-0.630** (0.283)	-0.548** (0.245)	-0.020** (0.009)		-0.925* (0.487)	-0.879* (0.444)	-0.032* (0.019)
Interaction		-2.251* (1.333)	-1.264* (0.712)	-0.057** (0.028)		-3.699 (2.214)	-2.065* (1.211)	-0.106** (0.045)
Observations	925,174	925,174	925,174	526,294	925,174	925,174	925,174	526,294
R^2	0.923	0.924	0.870	0.207	0.938	0.939	0.906	0.320
Mean Dependent Variable	9.51	9.51	5.23	0.50	17.12	17.12	10.23	0.80

Notes: Observations are at the county-month level, and are population weighted. Standard errors are clustered by state. All regressions include state, month, and year fixed effects. Additional controls include total population, population density, percent male, percent white, percent black, percent hispanic, percent unemployed, percent 18 and under, percent 65 and over, percent with a high school degree or less, percent poverty, and median income quintiles. Each column further controls for the size of the population for which suicide mortality is being measured (e.g. Column 2 additionally controls for the size of the black population). “Low Educ” is defined as having a high school degree or less.

Table 2.17: Prescriptive Authority and Days in Poor Mental Health

	(1) Full sample	(2) Black	(3) Middle Age	(4) Low Educ.	(5) Low Inc.	(6) Male	(7) Female
Indep. Prescriptive Authority	-0.093 (0.083)	-0.007 (0.105)	-0.244** (0.101)	-0.095 (0.136)	-0.109 (0.202)	-0.116 (0.078)	-0.068 (0.111)
Low Psychiatrist-to-Pop. Ratio	0.113 (0.075)	0.158 (0.180)	-0.047 (0.055)	0.157 (0.124)	0.201** (0.097)	0.096 (0.074)	0.130 (0.080)
Interaction	-0.186* (0.109)	-0.705*** (0.203)	-0.032 (0.100)	-0.267* (0.140)	-0.240 (0.147)	-0.147 (0.105)	-0.219* (0.125)
Observations	5,670,468	448,457	1,080,687	2,296,282	1,971,688	2,233,330	3,437,138
R^2	0.044	0.043	0.075	0.044	0.054	0.041	0.040
Mean Dependent Variable	3.33	3.77	3.66	3.86	4.34	2.77	3.85

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. Additional controls include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles are also included. “Low Educ” is defined as having a high school degree, less than a high school degree, or missing education information. “Low Inc” is defined as being in the bottom two quintiles of income.

Table 2.18: Prescriptive Authority and 21+ Days in Poor Mental Health

	(1) Full sample	(2) Black	(3) Middle Age	(4) Low Educ.	(5) Low Inc.	(6) Male	(7) Female
Indep. Prescriptive Authority	-0.002 (0.002)	0.002 (0.003)	-0.008*** (0.003)	-0.002 (0.004)	-0.003 (0.005)	-0.002 (0.002)	-0.002 (0.003)
Low Psychiatrist-to-Pop. Ratio	0.002** (0.001)	0.007* (0.004)	-0.003*** (0.001)	0.004* (0.002)	0.005** (0.002)	0.003** (0.001)	0.002 (0.001)
Interaction	-0.003 (0.002)	-0.020*** (0.005)	0.006*** (0.002)	-0.005** (0.002)	-0.004 (0.003)	-0.003 (0.002)	-0.003 (0.002)
Observations	5,670,468	448,457	1,080,687	2,296,282	1,971,688	2,233,330	3,437,138
R^2	0.024	0.022	0.049	0.025	0.031	0.024	0.023
Mean Dependent Variable	0.06	0.07	0.07	0.07	0.08	0.05	0.06

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. Additional controls include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles are also included. “Low Educ” is defined as having a high school degree, less than a high school degree, or missing education information. “Low Inc” is defined as being in the bottom two quintiles of income.

Table 2.19: Self-Reported Mental Health: Full Regression Results

	Days in Poor Mental Health			21+ Days in Poor Mental Health		
	(1) Full sample	(2) Full sample	(3) Low Educ.	(4) Full sample	(5) Full sample	(6) Low Educ.
Indep. Prescriptive Authority	-0.134* (0.072)	-0.093 (0.083)	-0.095 (0.136)	-0.003 (0.002)	-0.002 (0.002)	-0.002 (0.004)
Low Psychiatrist-to-Pop. Ratio		0.113 (0.075)	0.157 (0.124)		0.002** (0.001)	0.004* (0.002)
Interaction		-0.186* (0.109)	-0.267* (0.140)		-0.003 (0.002)	-0.005** (0.002)
Male	-0.940*** (0.028)	-0.940*** (0.028)	-0.917*** (0.060)	-0.012*** (0.001)	-0.012*** (0.001)	-0.013*** (0.002)
White	0.123 (0.104)	0.122 (0.104)	-0.101 (0.172)	-0.002 (0.003)	-0.002 (0.003)	-0.007* (0.004)
Black	-0.087 (0.175)	-0.087 (0.175)	-0.513* (0.257)	-0.005 (0.005)	-0.005 (0.005)	-0.017*** (0.006)
Hispanic	-0.664*** (0.139)	-0.663*** (0.139)	-1.287*** (0.200)	-0.019*** (0.004)	-0.019*** (0.004)	-0.035*** (0.005)
Employed	-1.246*** (0.040)	-1.246*** (0.040)	-1.669*** (0.062)	-0.033*** (0.001)	-0.033*** (0.001)	-0.045*** (0.002)
Health Insurance	-0.355*** (0.047)	-0.355*** (0.047)	-0.103* (0.055)	-0.009*** (0.001)	-0.009*** (0.001)	-0.004*** (0.001)
Age: 18 to 34	1.444*** (0.064)	1.445*** (0.064)	1.877*** (0.114)	0.010*** (0.001)	0.010*** (0.001)	0.023*** (0.004)
Age: 35 to 44	1.651*** (0.072)	1.652*** (0.072)	2.255*** (0.119)	0.025*** (0.002)	0.025*** (0.002)	0.042*** (0.004)
Age: 45 to 54	1.667*** (0.057)	1.667*** (0.057)	2.322*** (0.108)	0.032*** (0.001)	0.032*** (0.001)	0.050*** (0.003)
Age: 55 to 64	0.760*** (0.042)	0.760*** (0.043)	1.149*** (0.115)	0.015*** (0.002)	0.015*** (0.002)	0.024*** (0.004)
Age: 65 and Over	-1.225*** (0.046)	-1.225*** (0.046)	-1.240*** (0.130)	-0.029*** (0.002)	-0.029*** (0.002)	-0.030*** (0.005)
Education: HS or Less	0.451*** (0.088)	0.451*** (0.087)		0.005 (0.003)	0.005 (0.003)	
Education: Some College or More	0.059 (0.085)	0.059 (0.084)		-0.007** (0.003)	-0.007** (0.003)	
Income: 1st Quintile	1.973*** (0.053)	1.973*** (0.053)	1.652*** (0.083)	0.043*** (0.002)	0.043*** (0.002)	0.037*** (0.003)
Income: 2nd Quintile	0.665*** (0.022)	0.665*** (0.022)	0.268*** (0.032)	0.010*** (0.001)	0.010*** (0.001)	0.001 (0.001)
Income: 3rd Quintile	0.044 (0.049)	0.044 (0.049)	-0.397*** (0.061)	-0.005*** (0.001)	-0.005*** (0.001)	-0.016*** (0.002)
Income: 4th Quintile	-0.467*** (0.051)	-0.467*** (0.050)	-0.873*** (0.066)	-0.017*** (0.001)	-0.017*** (0.001)	-0.028*** (0.002)
Income: 5th Quintile	-1.274*** (0.065)	-1.274*** (0.065)	-1.633*** (0.099)	-0.035*** (0.002)	-0.035*** (0.002)	-0.044*** (0.003)
Observations	5,670,468	5,670,468	2,296,282	5,670,468	5,670,468	2,296,282
R ²	0.044	0.044	0.044	0.024	0.024	0.025
Mean Dependent Variable	3.33	3.33	3.86	0.06	0.06	0.07

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. “Low Educ” is defined as having a high school degree, less than a high school degree, or missing education information. For both income and education, the omitted category is a dummy equal to one if the relevant information is missing and zero otherwise.

Table 2.20: Self-Reported Number of Days in Poor Mental Health: Robustness (Leave Out)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ID	ID	NV	UT	CO	WY	NE	NM	ND	MI
Indep. Prescriptive Authority	-0.111 (0.080)	-0.078 (0.087)	-0.104 (0.083)	-0.035 (0.094)	-0.093 (0.083)	-0.100 (0.083)	-0.095 (0.083)	-0.096 (0.083)	-0.091 (0.085)
Low Psychiatrist-to-Pop. Ratio	0.117 (0.075)	0.114 (0.075)	0.115 (0.075)	0.116 (0.075)	0.113 (0.075)	0.115 (0.075)	0.114 (0.075)	0.114 (0.075)	0.151** (0.068)
Interaction	-0.205* (0.113)	-0.191* (0.109)	-0.140 (0.123)	-0.202* (0.110)	-0.185* (0.109)	-0.193* (0.112)	-0.171 (0.125)	-0.189* (0.111)	-0.221** (0.109)
Observations	5,571,593	5,608,071	5,554,971	5,543,660	5,585,488	5,499,136	5,566,294	5,602,195	5,552,622
R ²	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IA	IA	WI	KY	MD	DE	RI	CT	ME	HI
Indep. Prescriptive Authority	-0.095 (0.083)	-0.123 (0.083)	-0.093 (0.084)	-0.081 (0.109)	-0.091 (0.084)	-0.094 (0.084)	-0.091 (0.083)	-0.132* (0.070)	-0.084 (0.088)
Low Psychiatrist-to-Pop. Ratio	0.114 (0.076)	0.117 (0.075)	0.113 (0.075)	0.113 (0.074)	0.114 (0.075)	0.113 (0.075)	0.113 (0.075)	0.112 (0.075)	0.114 (0.075)
Interaction	-0.187* (0.110)	-0.094 (0.129)	-0.189* (0.110)	-0.185 (0.112)	-0.187* (0.110)	-0.185* (0.109)	-0.185* (0.110)	-0.176 (0.111)	-0.187* (0.110)
Observations	5,571,881	5,591,738	5,534,032	5,534,900	5,600,670	5,588,177	5,568,789	5,579,304	5,573,140
R ²	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. Additional controls include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles are also included. "Low Educ" is defined as having a high school degree, less than a high school degree, or missing education information. "Low Inc" is defined as being in the bottom two quintiles of income.

Table 2.21: Self-Reported 21+ Days in Poor Mental Health: Robustness (Leave Out)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ID	ID	NV	UT	CO	WY	NE	NM	ND	MI
Indep. Prescriptive Authority	-0.003 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.001 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)
Low Psychiatrist-to-Pop. Ratio	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.003** (0.001)
Interaction	-0.003 (0.002)	-0.003 (0.002)	-0.001 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.003 (0.002)	-0.003 (0.002)
Observations	5,571,593	5,608,071	5,554,971	5,543,660	5,585,488	5,499,136	5,566,294	5,602,195	5,552,622
R^2	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IA	IA	WI	KY	MD	DE	RI	CT	ME	HI
Indep. Prescriptive Authority	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.003* (0.001)	-0.002 (0.002)
Low Psychiatrist-to-Pop. Ratio	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)
Interaction	-0.003 (0.002)	-0.002 (0.002)	-0.003 (0.002)	-0.002 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.002 (0.002)	-0.003 (0.002)
Observations	5,571,881	5,591,738	5,534,032	5,534,900	5,600,670	5,588,177	5,568,789	5,579,304	5,573,140
R^2	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects. Additional controls include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles are also included. "Low Educ" is defined as having a high school degree, less than a high school degree, or missing education information. "Low Inc" is defined as being in the bottom two quintiles of income.

Table 2.22: Self-Reported Mental Health: Including State-Specific Linear Time Trends

	Days in Poor Mental Health			21+ Days in Poor Mental Health			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full	Full	Low Educ.	(mean)	Full	Full	Low Educ.
	sample	sample		low_educ	sample	sample	
Indep. Prescriptive Authority	-0.016	-0.032	0.025		0.000	-0.000	-0.000
	(0.159)	(0.162)	(0.180)		(0.002)	(0.002)	(0.002)
Low Psychiatrist-to-Pop. Ratio		-0.059	-0.103			-0.001	-0.002
		(0.067)	(0.097)			(0.001)	(0.002)
Interaction		0.088	0.102			0.002	0.002
		(0.123)	(0.144)			(0.002)	(0.002)
Observations	5,670,468	5,670,468	2,296,282		5,670,468	5,670,468	2,296,282
R^2	0.045	0.045	0.045		0.025	0.025	0.025
Mean Dependent Variable	3.33	3.33		3.86	0.06	0.06	

Notes: Observations are at the individual level with sample weights. Standard errors are clustered by state. All regressions include state and year fixed effects, as well as state-specific linear time trends. Additional controls include dummies for whether the respondent is male, white, black, hispanic, employed, and has health insurance. Dummies for age groups, education groups, and income quintiles are also included. “Low Educ” is defined as having a high school degree, less than a high school degree, or missing education information. “Low Inc” is defined as being in the bottom two quintiles of income.

Chapter 3

Does Physician Pay Affect Procedure Choice and Patient Health? Evidence from Medicaid C-section Use*

3.1 Introduction

Despite a growing literature that explores how doctors respond to financial incentives over procedure choice Coey (2013); Grant (2009); Gruber and Owings (1996); Gruber et al. (1999); Hadley et al. (2001); Keeler and Fok (1996); Yip (1998), few have linked changes in pay to health outcomes. In one of the most credibly identified recent papers looking at the impact of payment on medical treatment, Clemens and Gottlieb (2014) estimate small health impacts of changing financial incentives for doctors, “albeit with limited precision”. The lack of evidence on the effect of payment policy on health outcomes is an important gap in this literature, especially as changing doctor payment is often suggested as a way to decrease healthcare costs.

In this paper, I look at the effects of changing doctor payment on procedure choice and health outcomes for Medicaid patients. In particular, I study how changes in Medicaid reimbursement rates for birth procedures affect C-section use and infant health. C-sections are a useful case study for how payment can affect procedure choice, as doctors have a lot of discretion over C-section

*Materials from this chapter were presented at the Annual Meetings of the Population Association of America (2014). I am grateful to Janet Currie, Ilyana Kuziemko, Tom Vogl, Tom Winberry, and the seminar participants at Princeton for many helpful comments. This research received financial support from the Griswold Center for Economic Policy Studies.

use. In addition, as the most commonly performed operating room procedure in the US, C-sections are intrinsically important for costs Healthcare Cost and Utilization Project (2012).

As in Gruber et al. (1999), I find that doctors' perform more C-sections when the the "pay differential"—how much more doctors are paid to perform a C-section than a vaginal birth—increases. However, increases in the Medicaid pay difference are also associated with improvements in infant health. While previous work on high income mothers has found health benefits associated with lowering the use of C-sections Johnson and Rehavi (2013), my results suggest that C-section rates among women on Medicaid may actually be too low. This interpretation is consistent with the fact that the primary C-section rate for women on Medicaid is approximately 30 percent lower than that of women covered by private insurance, despite these women being in worse average health.¹ Policies targeting "unnecessary procedures" in the Medicaid population may not only be misplaced, but harmful if Medicaid enrollees on the margin have a higher benefit for the procedure than the population average.

My empirical strategy ties changes in state-level Medicaid reimbursement rates for birth procedures over time to C-section use and infant death rates. As Medicaid payment rates are not compiled centrally, previous work using Medicaid reimbursement rates has relied upon secondary sources for a small sample of states and years. However, a long panel is necessary to detect effects on a low frequency outcome such as infant death. Therefore, one contribution of this paper is the collection of reliable historical rate information from each state's Medicaid office.^{2,3} While Medicaid pays both doctors and hospitals each time a C-section is performed, Foo et al. (2015) find that the rate paid to the doctor primarily determines procedure choice. Thus, in this paper I abstract from Medicaid payments to hospitals.

The Medicaid C-section pay differentials are not correlated with average population health or local economic conditions, and are plausibly exogenous to the question of C-section use and infant

¹The primary C-section rate refers to C-sections performed on women who have not had a previous C-section.

²A complete list of which states and years are used is reported in Appendix Table 3.14. Not all states were able to give out historical reimbursement rate data, and some states did not have records going all the way back to the start of the sample period.

³These data are posted on my website: <http://scholar.princeton.edu/daalexand>

health. To further support the argument that changes in doctor pay drive procedure use, I show that the Medicaid pay difference is not associated with procedure use in counties where doctors are paid an annual salary. Unlike most doctors who are paid separately for each service provided, those in military and Indian Health Service hospitals are salaried, and thus provide a useful placebo group.

An increase in the Medicaid pay difference of \$100 is associated with a four percent increase increase in the probability of a C-section for women on Medicaid, confirming Gruber et al. (1999) on a much larger sample of states and years. This association disappears in counties with a salaried hospital, as expected if pay differences themselves (and not unobservable difference in patient health that may be correlated with pay difference) drive procedure choice. Similarly, the Medicaid pay difference does not affect C-section use among privately insured mothers. Finally, I show that the pay difference has no effect on procedure use for women who have had a previous C-section—cases where medical guidelines limit doctor discretion. The association between the pay difference and C-section use is unique to women on Medicaid with no previous C-section, whose doctors actually face the pay differential when making decisions.

Infant death rates, on the other hand, go down when Medicaid pays relatively more for C-sections. The positive association between the pay differential and infant health is largest for women on Medicaid with ex ante high-risk pregnancies—most pregnancies where C-sections are medically indicated.⁴ Again, there is no association between infant death rates and the pay difference for Medicaid births in counties with a salaried hospital, nor for the privately insured. Unfortunately, I have no information on the health of the mother, so I do not see the complete picture of how changes in incentives to perform C-sections affect health.⁵ What I can say is that when doctors are paid relatively more for C-sections, these mothers receive more C-sections and infant health improves.

⁴Ex ante high-risk pregnancies refer to those with risk factors listed on the birth certificate, where the risk factor would have been known prior to birth. More information is provided in the appendix.

⁵While maternal mortality rates can be constructed using vital statistics mortality files, maternal death in childbirth in the US is extremely rare. The more salient health concern for mothers considering a C-section is postnatal morbidity. As I do not know a good way to measure childbirth related morbidity, I leave this question to future research.

While there is evidence that some C-sections performed in the US are unnecessary Johnson and Rehavi (2013); Currie and MacLeod (2013, 2008), and it is true that Medicaid pays for almost half of all US births (see Figure 3.1), it does not follow that dis-incentivizing the use of C-sections for women on Medicaid will either save money or improve health outcomes. If C-section rates are too low for women on Medicaid, policies that reduce financial incentives to perform C-sections on these mothers could actually harm infant health.⁶ More generally, my results suggest that procedures which are overused in the general population are not necessarily overused for all groups. Trying to save money for Medicaid by discouraging the use of these types of “discretionary” procedures could backfire if they are actually underutilized for low-income patients.

The rest of the paper is organized as follows. Sections 3.2 and 3.3 lay out the conceptual framework and the data, respectively. Section 3.4 provides medical background on when a C-section should be performed. Section 3.5 describes the empirical strategy, which makes use of state-time variation in the Medicaid pay differential. Results are provided in Section 3.6. For Medicaid mothers seen by fee-for-service doctors, the pay difference is positively correlated with the probability of receiving a C-section, and is negatively correlated with infant death. In order to see which mothers benefit from high pay differentials, I differentiate between women with low risk and high risk pregnancies. Using this distinction, I examine the associations between the pay differential and infant outcomes: infant death, birth weight, gestational age, and the number of prenatal visits. Section 3.7 concludes.

3.2 Conceptual Framework

The intuition of this paper is straightforward—when doctors choose whether or not to perform a C-section, the relevant price is how much more they are paid for a C-section compared to a vaginal birth. To further establish the relationship between the pay difference and C-section use for fee-

⁶The Southern Legislative Council, for example, argues that Medicaid programs can save a substantial amount of money through policies that discourage unnecessary C-sections. The Council specifically mentions equalizing Medicaid payments for C-sections and vaginal birth as a popular and lucrative solution. (https://www.slcatlanta.org/QoM/qom.php?post_id=130)

for-service doctors, I compare them to doctors paid a flat salary, who would not be expected to respond to the pay difference.

A formal model is included in the Appendix, which is a simple extension of that in Gruber et al. (1999). In the model, the pay difference is the relevant variable influencing procedure choice, and under a salaried regime, this pay difference is zero. If many doctors are performing C-sections in marginal cases because they are more profitable than the alternative, then either a decrease in the pay difference or being salaried should lead to lower use of the procedure.⁷ The model predicts that the probability of C-section is increasing in the pay difference in fee-for-service hospitals, while in salaried hospitals it is constant over the pay differential.

If all infants born in counties with a salaried hospital were in fact born at that hospital, then I could run the regression

$$\begin{aligned}
 C - section_{ict} = & \beta_0 + \beta_1 Pay\ Difference_{st} \\
 & + \beta_2 Pay\ Difference_{st} * Salaried\ Hosp_{ct} \\
 & + \beta_3 Salaried\ Hospital_{ct} + \epsilon_{ict},
 \end{aligned} \tag{3.1}$$

where i indicates that the variable is defined at the individual level, c at the county level, s at the state level, and t that the variable is time varying. I would expect β_1 to be positive—doctors at fee-for-service hospitals increase procedure use in response to increases in the relative reimbursement rate. Likewise, I would predict that the fee difference would have no effect on doctors in salaried counties: $\beta_1 + \beta_2 = 0$. Finally, the model implies that all else being equal, the procedure choices of salaried doctors and fee-for-service doctors will be the same: $\beta_3 = 0$.

In the real world, however, many counties with a salaried hospital also have at least one other

⁷The statement that being salaried should lead to lower procedure use assumes that the substitution effect dominates the income effect. Conversely, if the income effect dominates, the fact that doctors are more wealthy due to the increased pay difference will lead to a decrease in C-sections; inducement will go down. Previous empirical studies have found that increasing relative reimbursement for intensive procedures increases their use Gruber et al. (1999); Hadley et al. (2001); Grant (2009). I follow these studies and assume that the substitution effect does in fact dominate: an increase in the pay difference will lead to an increase in the use of C-sections.

hospital. According to survey data from the American Hospital Association, in the year 2000, on average 43 percent of births in counties with a salaried hospital occur in that hospital. The fact that I do not know the exact hospital of birth introduces noise, which weakens the physician level predictions. I still expect that $\beta_1 > 0$, $\beta_2 < 0$, and $\beta_3 = 0$. In the data, however, β_1 and β_2 may not exactly offset, as some infants born in counties with a salaried hospital are born elsewhere.

3.3 Data

Studying the link between physician pay, procedure choice, and patient health is difficult primarily because of the scarcity of data. To remedy this problem, I combine data from several sources: vital statistics data on births and infant deaths, survey data about hospitals, and newly collected data on Medicaid pay rates. As Medicaid is administered at the state level, it provides a rich source of variation in physician pay. Unfortunately, data on Medicaid pay rates are not collected centrally. One contribution of this paper, therefore, is the compilation of reliable rate information. Unlike the Medicaid payment data used in other papers, which are built up from a combination of secondary sources, the payment data used in this paper is constructed from primary sources obtained by reaching out to each state's Medicaid office.⁸ I use these new data to build upon previous work, which was not able to look at health effects.

3.3.1 Vital Statistics Data

The primary data sources are the National Vital Statistics System's linked birth and infant death records, which record all infant births and deaths in the US, from 1990-1991 and 1995-2008 (the linked files were not created in 1992-1994). All outcome variables used in this analysis originate in the vital statistics data: C-section use, infant and pregnancy characteristics, and infant death.

⁸Gruber et al. (1999) used HCUP data from nine states (CA, CO, FL, IL, IA, MA, NJ, WA) from 1988–1992. Their data on payment came from secondary sources stitched together: for 1988 and 1992, they used data from by the American College of Obstetricians and Gynecologists (ACOG); for 1989 and to supplement missing values in 1988, the Physician Payment Review Commission; for 1990, Holahan (1993) supplemented with information from ACOG; for 1991, from Singh et al. (1993).

In addition, the natality data include county and month of birth, as well as detailed demographic and medical risk factors associated with the mother, infant, and birth event. These variables allow me to control for individual characteristics which may be correlated with place of birth and birth procedure.

The main analysis sample includes only women who have not had a previous C-section—whose doctors are choosing between a first C-section or a vaginal birth (these women could have had other births, as long as they never had a C-section). This sample restriction is important, as doctors have much more discretion over procedure choice when contemplating a first C-section. Rates of vaginal birth after Cesarean (VBAC) are very low in the US (MacDorman et al. (2011) reports that in 2007, the US VBAC rate was just 8.3 percent), and doctors have long followed the rule of thumb that “once a Cesarean, always a Cesarean”. This recommendation has been heavily criticized in recent years, but was closely followed over much of the sample period. In addition, because of the high risk of serious complications for women who have undergone a previous C-section, it is unlikely that the Medicaid pay differential would influence a doctor’s decision in the case of repeat C-sections. For all of these reasons, I expect the pay difference to impact physician behavior for primary C-sections only, and will use mothers who had previous C-sections only in placebo tests.

As the identification strategy uses Medicaid pay differentials, it is important to determine which mothers are covered by Medicaid at the time of birth. Unfortunately, the vital statistics data do not include insurer over the study period. However, insurance coverage information was added in the most recent natality files for some states. I use these new data to predict Medicaid coverage for my sample of mothers. This prediction is quite good; the pseudo R^2 is 0.300 (details on this exercise can be found in the data appendix). I call mothers whose predicted probability of Medicaid coverage are in the top third “high predicted Medicaid mothers”, and use this group as my preferred sample. “Low predicted Medicaid mothers” are those whose predicted probability of Medicaid coverage are in the bottom third, and are used throughout the analysis as a placebo group.

The main health measure used to study how C-section use affects infant health is infant survival, as death is an unambiguously bad outcome. My preferred measure is whether the infant survived to the first birthday, though I also look at the probability of death over different time horizons. Birth weight is also often used as a measure of infant health; low birth weight (less than 2500 grams) is associated with a range of poor outcomes, both health related and economic, such as schooling and earnings Black et al. (2007). As the pay differential impacts only the type of birth and not the overall health of the infant pre-birth, however, there should be no effect on birth weight—except for a mechanical decrease due to C-sections being scheduled slightly before vaginal birth would occur, or an increase due to fluid in the lungs, which is expelled during a vaginal delivery but not a C-section.⁹

Finally, the sample is restricted to singleton births that occurred inside a hospital, in the 50 states and Washington, DC. Territories and protectorates of the US have very different public health insurance programs, and the incentives studied do not apply. Multiple births are also dropped from the analysis, as they are much more likely to develop complications that require delivery via Cesarean section (the C-section rate for twins was 75% in 2008) Lee et al. (2011).

3.3.2 Physician Reimbursement Data

The second data source contains Medicaid reimbursement rates for obstetric procedures at the state-month level, from 1990-2008 (the exact date range depends on the state). As Medicaid reimbursement rates are not collected centrally, the rate data was obtained directly from each state's Medicaid office. While some states were unable to provide this information, I was able to collect data on 36 states plus Washington, D.C., which cover approximately 85 percent of the population (details on which states and years are used are provided in Appendix Table 3.14).¹⁰ Over the study period, states varied both in reimbursement levels, as well as in how frequently the rates were

⁹Approximately 35 mL of fluid is expelled from an infant's lungs during a vaginal delivery. In the first few hours of life, infants born by C-section at term have overall lung volumes similar to those born by vaginal delivery, but the liquid volume is increased and the gaseous component decreased Milner et al. (1978).

¹⁰The raw data on reimbursement rates is available for download on my website: <http://scholar.princeton.edu/daalexand>

changed; some states' payments change at least once a year, and others only change once or twice over the sample period (see Table 3.1 for means).

As discussed in Section 3.2, the relevant choice variable for physicians deciding to perform a C-section is the pay differential—how much more they are paid for a C-section than a vaginal birth. This variable is constructed as the difference between amounts paid to doctors by Medicaid for these two procedures (details of the construction of the pay differential can be found in the data appendix).¹¹ Figure 3.2 plots the Medicaid pay differential for each state, over time. Payment schedules vary widely from state to state, with the rates and the pay difference both increasing and decreasing over the eighteen year period. On January 1st, 2000, reimbursement rates for a C-section with postpartum care ranged from less than \$600 or less in California, New Jersey, and Michigan to \$1,200 or more in Georgia, Arizona, Nevada, and Alaska. On the same date, the payment difference ranged from \$0 in Michigan and Montana to \$370.80 in Minnesota and \$410.72 in Nebraska. Table 3.1 reports summary statistics of these payment variables across states and time periods. My maintained assumption is that the timing of changes in these pay rates—and especially the differences between them—reflect the idiosyncratic politics of each state, rather than the preferences of legislators over C-section use.

Finally, an important institutional detail of the Medicaid program is the role of Medicaid Managed Care. States have flexibility not just with respect to Medicaid generosity, but also in how to administer the benefit. Over the past few decades, an increasing number of states have opted to contract with health maintenance organizations (HMOs) and other managed care organizations (MCOs) to provide care for part or all of their Medicaid recipients Duggan and Hayford (2013). In states and time periods where a large fraction of Medicaid recipients are covered by managed care organizations, the details of the Medicaid fee schedule—which mainly applies to traditional fee-for-service Medicaid—should have less of an impact.¹² To verify that the effect of pay difference

¹¹ As a placebo test, I also look at women with previous C-sections; in these regressions, the pay difference is between C-sections and vaginal birth after Cesarean (VBAC).

¹² One large category of managed care also uses the Medicaid fee schedule to reimburse doctors: Primary Care Case Management (PCCM). In PCCM, a primary care provider is paid a small monthly case management fee in return for monitoring the care of Medicaid beneficiaries. This fee is in addition to fee-for-service reimbursement for treatment Sekunda et al. (2001).

on procedure choice is stronger when MCOs are less active, I will examine how the impact of the pay differential on C-section use varies with respect to state-level use of managed care.

3.3.3 Hospital Data

The final data source is the American Hospital Association (AHA) annual survey, which contains information on whether counties contain government-run hospitals. The key distinction is between hospitals controlled by the federal government and other hospitals, because doctors employed by the government are salaried, while most others are paid fee-for-service.^{13,14} In contrast to the doctors working under a fee-for-service regime, the salary structure of doctors employed by the government is based on promotions and years of service, rather than the number of procedures performed (for more information on the pay of federally employed doctors, see Appendix Table 3.15). The AHA data let me use salaried doctors as a comparison group when looking at the effect of the pay differential on C-section use. The federally controlled hospitals with non-zero births are Air Force, Army, Navy and Indian Health Service hospitals—hospitals with no births recorded are not counted (Figure 3.3 shows the locations of these hospitals).

The Medicaid pay differential described in the previous section only applies to mothers covered by Medicaid. Doctors treating mothers covered by private insurance face a different pay difference, which I cannot observe. Therefore, I need to be able to target my analysis towards women covered by Medicaid at the time of birth. In order to focus on the Medicaid population, I use two complementary approaches. As described previously, I use recent data on Medicaid coverage at birth to predict the likelihood of Medicaid coverage for the full sample. In addition, I construct an index of Medicaid intensity at the county level using American Hospital Association data on the number

¹³It is difficult to find direct evidence on the prevalence of fee-for-service compensation, but the Community Tracking Survey (2008) supports this characterization. Among obstetricians/gynecologists who admitted a patient to a hospital in the last year, 73% report either being paid for performance (mostly based on factors reflecting “own productivity”), for their share of practice billings, or are solo practitioners. Additionally, 63% are independent contractors or full or part owners of their practice. While the sample size is small (N=284), the survey supports the importance of fee-for-service reimbursements for the typical obstetrician/gynecologist.

¹⁴Kaiser Permanente, a large integrated managed care consortium, also uses a salaried incentive scheme rather than fee-for-service. However, all Kaiser hospitals are located in densely populated counties where they account for only a very small fraction of births per year, and thus cannot be studied using the framework of this paper.

of Medicaid discharges per hospital.

The county-level “Medicaid intensity” index is constructed by dividing the number of Medicaid discharges by the total number of hospital beds, and averaging the ratio over the sample period. I expect that doctors with many Medicaid patients should be more sensitive to the pay difference than doctors who primarily deliver privately insured mothers. Alternatively, this index may identify counties where there is less error in the Medicaid prediction equation. Either way, the index captures the geographic concentration of Medicaid recipients, and is used to identify counties where the Medicaid pay differential is predicted to have an especially large impact.

The hospital level is collapsed to the county-year level, and merged into the vital statistics files. In the combined data, whether a federally controlled (salaried) hospital existed in the county and year of birth is known for all births in the sample. A drawback to this strategy is that the exact hospital of birth for infants born in counties with more than one hospital is not known. I am limited by the data to studying the impact of having a salaried hospital in the county and year of birth, rather than being born in a salaried hospital. This data limitation, however, should bias me against finding a difference between counties which do and do not contain a salaried hospital.

3.4 When Should a C-section Be Performed?

C-sections are generally used when pregnancy complications make traditional vaginal birth difficult, or when vaginal birth puts the mother or child in danger. While not necessary in the majority of deliveries, C-sections are lifesaving for both mother and infant in many situations. The most common reasons given for primary C-section in the US are arrest of labor, abnormal or indeterminate fetal heart rate tracing, fetal malpresentation, multiple gestation, and suspected fetal macrosomia; over half of all C-sections are indicated by the top two categories: arrest of labor, and abnormal fetal heart rate tracing Barber et al. (2011).

The primary C-section rate of women predicted to be on Medicaid is much lower than for women predicted to be privately insured. Should we expect fewer C-sections to be performed on

women on Medicaid, due to differences in medical conditions or demographics? Birth certificates do not record the most common indications for C-section in the US: arrest of labor and abnormal fetal heart rate. However, many health and behavioral characteristics recorded on the birth certificate are related to C-section use (Table 3.2, Panel A).

Women predicted to be covered by Medicaid in my sample tend to have riskier and less healthy pregnancies along many dimensions, as is shown in Table 3.2. Pregnancies of women predicted to be on Medicaid display higher rates of fetal distress (meconium staining) than those predicted to be covered by private insurance. In addition, these women are over four times as likely to smoke during pregnancy than those predicted to be privately insured. Smoking is associated with problems involving the placenta, such as an increased risk of placenta previa, a condition requiring C-section. Not surprisingly, women predicted to be on Medicaid are also more likely to have acute or chronic lung disease; pulmonary conditions, if poorly controlled, can also adversely affect pregnancy. On the other hand, women predicted to be on Medicaid are younger than their privately insured counterparts. Hence, they have lower incidence of medical risk factors associated with age, such as chronic hypertension, diabetes, and breech birth.

Using 2012 natality data which reports the insurer, I find that the primary C-section rate of women on Medicaid is three to four percentage points (approximately fifteen percent) lower than that of privately insured women, depending on the specification. Table 3.3 reports regressions comparing C-section use between those recorded to be on Medicaid and private insurance, and then between Medicaid plus the self-insured and private insurance. Self-insured births are combined with Medicaid births in the second set of regressions, as many of these births are likely ultimately covered by Medicaid; the socioeconomic status of self-insured women is similar to that of those covered by Medicaid, and hospitals often sign up Medicaid eligible women at the time of birth Heberlein et al. (2011).¹⁵ As shown in Table 3.3, the gap in primary C-section use between those on Medicaid versus private insurance use remains nearly unchanged when controls are added for

¹⁵In 1986, presumptive eligibility in Medicaid was established as a state option to improve access to timely care for uninsured pregnant women (Section 1920 of the Social Security Act, 42 U.S.C. 1396r-1). Hospitals, doctors, and pharmacies are paid for the services they provide during this temporary coverage period, even if an individual does not complete the full application (Families USA, Families USA (2011))

demographic characteristics, medical risk factors, and state, month, and day of the week fixed effects. Thus, there does not appear to be a medical reason for the lower C-section rates among women covered by Medicaid than those that are privately insured.

While significant in both my main data (1990-2008) and the 2012 natality data, the gap in C-section use between women on Medicaid and the privately insured is larger in the earlier data. It could be true that the gap in C-section use between mothers on Medicaid and the privately insured has gotten smaller over time—perhaps as C-sections have become more common, doctors have become more likely to perform them on Medicaid patients. However, these two datasets are difficult to directly compare. For example, their sample composition is different; the states with data on the pay differential are a different group than those using the revised birth certificate in 2012, which are included in Table 3.3. In addition, the cut off in predicted probability of Medicaid coverage used to divide the main sample is fairly arbitrary. No matter how the data is cut, however, there is strong support for the fact that women on Medicaid are less likely to receive a C-section than those that are privately insured—and that this is not explained by differences in underlying health or demographic factors.

3.5 Empirical Strategy

A linear probability model is used to measure the impact of the Medicaid pay differential on C-section use and infant health. In the main specification, I compare the effect of changing pay differentials on outcomes in counties with and without salaried hospitals, for mothers predicted to be on Medicaid. For robustness, and to show that the results are not spurious, I repeat the analysis on mothers predicted to be privately insured. The main explanatory variables are the pay differential between C-section and vaginal birth, $Pay\ Difference_{st}$, an indicator for whether the county contained a salaried hospital at the time of birth, $Salaried\ Hospital_{ct}$, and their interaction $Pay\ Difference_{st} * Salaried\ Hosp_{ct}$. The outcome is first whether a C-section was performed, and later infant survival—during the neonatal period, which is most closely related to birth out-

comes, or to the standard one-year threshold.

$$\begin{aligned}
Outcome_{ict} = & \beta_0 + \beta_1 Pay\ Difference_{st} \\
& + \beta_2 Pay\ Difference_{st} * Salaried\ Hosp_{ct} \\
& + \beta_3 Salaried\ Hospital_{ct} + \beta_4 X_{ict} + \lambda_s + \gamma_{year} + \epsilon_{ict}
\end{aligned} \tag{3.2}$$

Here, i , c , and s indicate that the variables are defined at the individual, county, and state level, respectively, and t that the variable is time varying. Many individual characteristics are also controlled for in X_{ist} : the mother's age, race, marital status and education, health and delivery complications, birth parity, and the weekday of birth. In addition, full sets of state and year fixed effects are included, λ_s and γ_{year} , as unobserved factors impacting delivery and prenatal health may differ between states or over time. Finally, to allow for arbitrary correlation of errors within counties, the standard errors are clustered at the county level.¹⁶

After establishing the link between the pay differential, procedure choice, and infant health, I examine which births benefit from high pay differentials. I focus on births which occurred in fee-for-service hospitals, and differentiate between ex ante low and high risk pregnancies. As high risk pregnancies are where C-sections are usually indicated, finding a larger effect on these pregnancies would be consistent with a story where doctors respond to increased financial incentives for C-sections by performing them on women who were previously just below the C-section threshold. In this analysis, various infant health outcomes are regressed on the pay difference, a dummy for a high risk pregnancy, and their interaction, along with patient level covariates and state and year fixed effects. I run the following regression,

¹⁶The results are nearly identical if I instead cluster the standard errors at the state level. I choose to cluster at the county level because the number of states in my sample is 36 plus Washington D.C., which falls below the rule of thumb of 50 clusters being the approximate lower bound to avoid the problems associated with too few clusters Cameron and Miller (2015).

$$\begin{aligned}
Health\ Outcome_{ict} = & \beta_0 + \beta_1 Pay\ Difference_{st} \\
& + \beta_2 Pay\ Difference_{st} * high\ risk_{ict} \\
& + \beta_3 high\ risk + \beta_4 X_{ict} + \lambda_s + \gamma_{year} + \epsilon_{ict}.
\end{aligned} \tag{3.3}$$

where the notation and covariates in X_{ict} are the same as above. The health outcomes are indicators for infant death and neonatal death. In order to investigate any potential effect of the fee differential on access to care, I also look at the number of prenatal visits. I follow previous literature and define a pregnancy as high risk if at least one of risk factor is noted on the birth certificate. Due to changes in recording over time, the group of complications used differs slightly across the sample (details can be found in the data appendix). The results, however, are robust to only using the set of complications that I can observe in all years.

The richness of the vital statistics data lets me control for the fact that women giving birth in areas with particular Medicaid pay differences or types of hospitals may differ from the norm along other observable dimensions. As shown in 3.2, Panel B, women who give birth in counties with and without salaried hospitals are fairly similar along observable dimensions, with a few caveats. For example, mothers who give birth in counties with a salaried hospital are more likely to be Hispanic or Native American. The higher percentage of Native American mothers is to be expected, as some of the federal hospitals are run by the Indian Health Service. Mothers in counties with federal government-controlled hospitals are also slightly healthier, which may be explained by them being a bit younger on average. In the regression analysis, I control for all demographic and health characteristics reported in Table 3.2.

The above empirical strategy has two main limitations: neither the exact insurance coverage nor the hospital type at birth is known for the majority of the sample. I address the first issue by predicting Medicaid coverage using the most recent natality file, which includes the insurance status of the birth. The second weakness—that the exact hospital of birth in counties with more

than one hospital is not known—is more difficult to address. However, the lack of individual level data on hospital type will bias my results towards finding no difference between counties with and without salaried hospitals. Since I do find a difference between these groups, I believe this particular data limitation is less of a concern. However, I cannot disprove a story where counties with a salaried hospital are different along some unobservable dimension, which also causes doctors not to respond to the pay differential.

In both specifications, I differentiate between women who are predicted to be on Medicaid and those who are not, as well as two subsamples: counties with few Obstetrician-Gynecologists (Ob-Gyns) per capita, and counties with a high density of Medicaid coverage. I focus on Medicaid mothers as their pregnancies are “treated” with the pay difference; mothers predicted to have private coverage are always shown for comparison. Counties which are underserved by Ob-Gyns are used to examine the role of capacity constraints in the decision to perform a C-section. If capacity constraints play an important role, increases in C-section rates for Medicaid mothers would be accompanied by decreases in C-section rates for other mothers—and this offset effect would be the largest in areas where providers are scarce. If capacity constraints are not important, however, C-section rates for Medicaid mothers could increase without a corresponding decrease in C-section rates for other mothers. Finally, I look at counties with a high density Medicaid coverage. If having more Medicaid patients makes the pay difference more salient for doctors, the effect of the pay difference on procedure use should be largest in these counties. Alternatively, my ability to predict which mothers are on Medicaid may be better in counties with high Medicaid density; if true, this would also predict finding larger effects in these counties.

Following Gruber et al. (1999), I assume that changes in pay difference over time are exogenous. The exogeneity assumption seems reasonable, as changes in reimbursement rates likely come from state legislators making decisions unrelated to the specific procedure choice of C-sections versus vaginal births. And furthermore, the timing of adoption is idiosyncratic; for example, the timing of legislation is often constrained by features of the legislature. One potential concern is that states reduce Medicaid payment rates in order to balance budgets during recessions. If reim-

bursements were high during good times, and economic conditions are correlated with health or C-section preferences, the estimates could be biased.¹⁷ As the identification strategy hinges on the difference in pay between C-sections and vaginal birth, rather than the levels of reimbursement, this concern is less of an issue. Still, a relationship between payment reductions and economic conditions could be a problem if states cut Medicaid payment rates differentially for C-sections in response to the business cycle. My results, however, are robust to controlling for state-level unemployment, both contemporaneous and lagged.¹⁸ Therefore, the relationship between the fee differences and local economic conditions does not appear to be a serious concern. Another issue raised in Gruber et al. (1999) is that a deterioration in fetal or maternal health that gave rise to more C-sections could also cause state policymakers to increase the pay differential. They address this possibility by including variables which control for the riskiness of the birth in their regressions, and find that it does not affect their estimates. Following these authors, an indicator for high-risk pregnancies is included in all regressions.

3.6 Results

3.6.1 Impact of Pay Difference on Procedure Choice

Consistent with Gruber et al. (1999), I find a positive correlation between the Medicaid fee difference and C-section use. Figure 3.4 shows this relationship graphically for the state of Arkansas. Arkansas is an ideal state to demonstrate the relationship between the fee difference and the C-section rate, for two reasons. First, the fee difference changed only once over the time period, making the pre- and post- comparisons straightforward. Second, the Medicaid pay difference is very relevant in Arkansas, as all doctors treating Medicaid patients are reimbursed on a fee-for-

¹⁷Whether health is pro- or countercyclical is a debate that is outside the scope of this paper; for a discussion see Ruhm (2000) and Dehejia and Lleras-Muney (2004), among others.

¹⁸Not only are the main coefficients unchanged, but the coefficients on the unemployment variables are not statistically significant.

service basis.¹⁹ Figure 3.4 plots the residuals from the following regression of C-section use on year fixed effects for Arkansas,

$$C\text{-}section_{ict} = \beta_0 + \gamma_{year} + \epsilon_{ict} \quad (3.4)$$

as well as the fee differential. The regression residuals are shown, rather than the average C-section rate, because the secular trends in C-section use over the last two decades swamp the changes in procedure use associated with the pay differential. After the pay differential drops in April of 2004, there is a corresponding drop in the regression residuals – the model goes from over-predicting the probability of C-section to under-predicting it. Furthermore, this drop only appears for women predicted to be on Medicaid (on the left), and not for women predicted to be privately insured (on the right).

More specifically, fee-for-service doctors respond to high pay differences by increasing the use of C-sections. For all women who are predicted to be covered by Medicaid at birth, I find that the payment variables are the expected sign, but are not statistically significant (Table 3.4, Column 1). When I narrow the sample to counties with high density of Medicaid coverage, however, I find much larger and significant effects. I expect doctors to be particularly affected by the pay differential when a large fraction of their patients are insured by Medicaid, and this is born out by the data. As shown in Column 2, women on Medicaid in high Medicaid counties are 0.6 percentage points more likely to have a C-section when the pay differential increases by \$100. This translates to a four percent increase in the probability of C-section for these women. In addition, as can be seen in Column 4, there are no analogous results for women predicted not to be on Medicaid.

In counties with low Ob-Gyn to population ratios (Columns 3 and 6), the results are less precisely estimated, but similar to those found in high Medicaid counties—suggesting that there are no offsetting effects on privately insured mothers. There is no evidence that providers decrease the use of C-sections for privately insured mothers in order to increase them for Medicaid mothers.

¹⁹The only type of Medicaid Managed Care used for medical coverage in Arkansas is Primary Care Case Management (in place since 1994), which uses a fee-for-service payment structure. Risk-based capitation reimbursement is only used for non-emergency transportation.

This suggests that capacity constraints are not binding in the choice to perform a C-section, and cannot explain the lower baseline C-section rates for Medicaid women. The similarities in effect size between the low provider and high Medicaid counties is a pattern that will be repeated throughout all the results. While there is some overlap between these two groups, they are not the same. Over half the mothers in low provider counties do not reside in a “high Medicaid density” county. The Ob-Gyn to population ratio is negatively correlated with Medicaid density, however, which likely explains why the effect sizes are larger than average in underserved areas.

As predicted, the C-section rate of Medicaid women at salaried hospitals is not responsive to the pay differential. The lack of response in salaried hospitals shows up in Table 3.4 as a negative and roughly offsetting coefficient on the interaction term between the pay difference and being in a salaried hospital in Columns 1 - 3; the effect of the pay difference rate on salaried doctors is the sum of the coefficients on pay difference and the interaction of the pay difference with hospital type. Again, as can be seen in Columns 4 - 6, there are no analogous results for women predicted not to be on Medicaid. The lack of response of salaried doctors suggests that the results for fee-for-service doctors reflect a true effect of the pay differential on procedure choice.

When the pay difference between procedures is zero, patients of salaried and fee-for-service doctors are equally likely to receive a C-section. The coefficient on the indicator for a county having a federally controlled hospital in Columns 1 - 3 of Table 3.4 is not statistically different from zero. The lack of a main effect of having a salaried hospital on the probability C-section indicates that the regression model is successful in controlling for population differences between mothers living in counties with and without a salaried hospital.

As a placebo check, the above analysis is repeated on women who have had a previous C-section; doctors have less discretion over procedure choice in these pregnancies.²⁰ As expected, there is no effect of the pay differential on procedure choice for these women. These results are reported in Appendix Table 3.16. Not only does the effect of compensation structure on procedure choice disappear when the outcome variable is a repeat C-section, but no pattern emerges related to

²⁰In these regressions, I use the pay differential between a C-section and vaginal birth after Cesarean (VBAC).

prevalence of Medicaid coverage. This null result again supports the interpretation that Columns 1 - 3 of Table 3.4 reflect a causal effect of the pay differential on C-section use.

Appendix Table 3.18 shows that the effect of the pay difference on C-section use is higher in states with lower use of Medicaid Managed Care. In Appendix Table 3.18, I stratify the sample by the fraction of Medicaid users enrolled in Medicaid Managed Care plans which reimburse doctors using capitation, as opposed to fee-for-service. In a capitated system, providers are paid a set amount for each enrollee assigned to them, per period of time, whether or not that person seeks care. The sample is divided into two groups: low and high use of capitation—based on the average fraction of Medicaid enrollees in Managed Care plans using capitation from 2003 (the earliest year of data available). While it is impossible to know how doctors in these managed care plans are actually paid, to the extent that it is not based on the Medicaid fee schedule, I expect to find weaker associations between the pay differential and C-section use in the states with high use of capitation. As can be seen in Appendix Table 3.18, that is what I find, albeit with low precision. In addition, the appendix reports results of a permutation test where pay differential time series are randomly assigned to states (Appendix Figure 3.5), and adding state-level linear time trends (Appendix Tables 3.19 and 3.20). The main results pass both of these additional robustness tests.

The results in Table 3.4 demonstrate that decreasing the fee difference by a modest amount could have an important impact on C-section use. As the exact insurance coverage and hospital of birth are unknown, these results are likely attenuated due to measurement error. Despite the coarseness of the specification, the estimate that a \$100 increase in the fee difference is associated with a 4% increase in C-sections is non-negligible, and can be interpreted as a lower bound on the true effect.

3.6.2 Impact of Pay Difference on Infant Health

The above results suggest that policy makers can exert considerable influence over procedure choice by changing how doctors are paid. However, more information is needed to recommend such a policy. In the case of C-sections, as in the case of many procedures, the effect of reduc-

ing procedure use on health is theoretically ambiguous. Reducing C-section rates could improve health by eliminating medically unnecessary procedures; on the other hand, it could worsen health by reducing the rates below the optimum level.

In Table 3.5, I show that the increased C-section rates for Medicaid mothers associated with high pay differentials are actually beneficial to the health of their children. The regression specification is the same as in the previous section, except with infant death as the dependent variable. The pattern across Table 3.5 is the same as Table 3.4, but opposite in sign. There is no effect in the general population, but once again the results appear in counties with high Medicaid coverage and low provider ratios. Infant death decreases when the pay difference goes up; a \$100 increase in the pay differential is associated with a 0.0006 percentage point decrease in infant death—6% of the mean for women on Medicaid. For salaried doctors, an increase in the pay differential does not affect infant death, again shown by the offsetting positive coefficient on the interaction term. In Tables 3.6 and 3.7, I repeat the analysis looking separately at death in and after the neonatal period—which is seen as more closely tied to events surrounding birth than deaths occurring later in the first year. The results for neonatal death closely mirror those of infant death; the association between the pay differential and death after the neonatal period is both smaller and less precisely estimated.

The positive association between C-sections and infant health is somewhat surprising from the perspective of the medical and health economics literature Baicker et al. (2006), which suggests that more C-sections are being performed in the US than are medically necessary. However, pregnant women on Medicaid are very different than those covered by private insurance. By definition these mothers are worse off financially, and as can be seen in Table 3.4, they have a much lower baseline probability of C-section. The primary C-section rate for mothers predicted to be on Medicaid is just two thirds of the rate of mothers predicted to be privately insured. While the rate for “all mothers” may be too high, the rate for Medicaid mothers may be too low, especially given that they also tend to be in worse health. Taking both these factors into account, it is not surprising that more C-sections are associated with better outcomes for this particular population.

That higher pay differentials are associated with both more C-sections and better infant health for Medicaid patients should give pause to policymakers intent on using payment incentives to lower costs. Any reform using incentives to change physician behavior must consider possible downstream effects on patient health. At least in the case of C-section use in the Medicaid population, reducing C-section rates seems to come with significant health costs for infants.

3.6.3 Pathways between pay differential and infant health

Finally, I investigate whether high pay differentials help all Medicaid mothers, or whether the benefits accrue to a particular group. I separate pregnancies into those considered *ex ante* high risk, and those that are not. As women with high-risk pregnancies would be the most likely to benefit from physicians performing C-sections at higher rates, I expect to find larger positive effects of high pay differentials on infant health among these women. I now estimate Equation 3.3, where an indicator for a high risk pregnancy is interacted with the pay differential. For both clarity and brevity, counties with salaried hospitals are dropped from this part of the analysis.

The benefits associated with high pay differentials do appear to be concentrated among women with high risk pregnancies. For women with at least one risk factor, a \$100 increase in the pay differential is associated with a -0.0007 to -0.001 percentage point decrease in the probability of infant death (see Table 3.8, Columns 1 - 3). As hypothesized, the effect of the pay differential is significantly larger for women with at least one recorded risk factor than for women with none. As death during the neonatal period (first 28 days) is considered to directly reflect prenatal, intrapartum, and neonatal care, in Table 3.9 I replicate the analysis for neonatal death, and find similar results. For comparison, Table 3.10 shows that the pay differential has much smaller estimated effects on infant deaths after the first 28 days.

Higher pay differentials are for the most part not associated with changes in birthweight or gestational age, as seen in Tables 3.11 and 3.12. The pay difference seems to be negatively correlated with birth weight for low-risk pregnancies, and positively correlated with birth weight for high-risk pregnancies, though the effects are generally not statistically significant. As previously discussed,

C-section use could be mechanically related to birth weight; either positively due to retained liquid in the lungs, or negatively, due to shorter gestational age.²¹ Neither of these mechanical effects, however, seem to be very large.

There are at least two potential mechanisms linking increases in the pay difference to health outcomes: direct benefits of higher C-section rates, and indirect benefits derived from changes to the delivery of and access to healthcare. I find the first explanation to be more compelling, especially as primary C-section rates for Medicaid women are far below that of the privately insured. Still, it could also be the case that higher pay differences affect infant health indirectly—for example, by making Ob-Gyns more likely to take on Medicaid patients, which could affect the quality and quantity of care provided).²² If true, increases in infant health would be correlated with—but not directly caused by—the increases in C-section rates. If increasing the Medicaid pay difference causes providers to increase the supply of services to the poor, we might expect the infant health gains to be particularly large in underserved areas. However, this is not the case; the pay difference always has a larger effect on outcomes in high Medicaid counties compared to low provider counties. On the other hand, higher pay differentials are associated with a small increase in the number of prenatal visits (see Table 3.13). While potentially supporting an access story, the effect on prenatal visits is both small in magnitude, and is larger for low-risk women than high-risk women—so it cannot explain the decrease in infant death.

Regardless of the particular mechanism driving the positive association between pay differentials and infant health, the above results do not necessarily imply that more C-sections should be performed, even for Medicaid women with high risk pregnancies. In this study, I was unable to take into account the increased health risks C-sections may bring to the mother, which could outweigh the benefits to infant health. The results do imply, however, that policies which try to control costs by lowering the use of procedures may be too heavy handed. Even for C-sections,

²¹ Approximately 35 mL of amniotic fluid is pushed from infants' lungs during vaginal birth. While this fluid is eventually cleared over the first six to twelve hours of life, its presence would lead to higher measured birth weight, on average.

²²Currie et al. (1995), for example, found that increasing Medicaid physician fees relative to private fees makes Medicaid patients more attractive, and was associated with declines in the infant mortality rate.

a procedure widely considered overused, lowering the use of procedures by reducing financial incentives can come with important health costs. These results support the findings of Currie and MacLeod (2013)—that improvements in doctors’ ability to diagnose when C-sections are indicated is likely better for health outcomes than reducing procedure use across the board. Diagnostic improvements could result in better outcomes for both high- and low-risk pregnancies, while simple payment reforms could lead to worse health outcomes for some groups, as shown here.

3.7 Conclusion

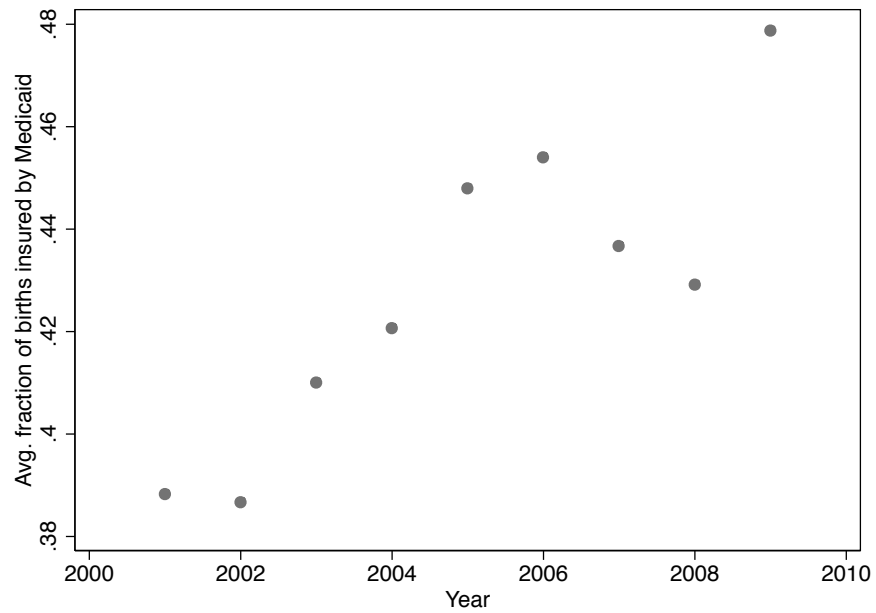
In hospitals where doctors are paid under a fee-for-service system, the more doctors are paid for a C-section relative a vaginal birth, the more likely they are to perform C-sections. In salaried hospitals, where the pay incentives to order more intensive procedures are eliminated, I find no effect of the pay difference on procedure use and lower C-section rates. This result alone could imply that many C-sections are due to the compensation structure, and a salaried system would eliminate perverse incentives and unnecessary surgeries. After demonstrating that the supply of C-sections is influenced by financial incentives, however, I find that higher pay differentials are associated not just with a higher probability of C-section, but with better health outcomes for high-risk Medicaid pregnancies. C-sections are widely considered to be overused—yet for marginal patients on Medicaid, there are measurable benefits for infant health from increasing their use.

In order to make a policy recommendation, however, it is necessary to have evidence on the effect of increasing C-section use on both maternal and infant health. Unfortunately, due to data limitations, I cannot look at how changing C-section use affects the health of Medicaid mothers. Furthermore, the main maternal health effects of C-sections lie in morbidity rather than mortality (maternal mortality rates in the US are very low), which is difficult to see in the data. Learning more about how C-section use affects maternal morbidity for women on Medicaid is a useful avenue for future research, and would allow for a cost-benefit analysis of changing C-section use for women on Medicaid.

Disincentivizing C-section use is associated with worse health for infants whose births are covered by Medicaid. This finding is an important indication that care must be taken when using financial incentives as a tool to change physician behavior. Reducing procedure use may have unintended consequences for patient health, which are likely to fall disproportionately on already vulnerable segments of the population. These potential side effects of cost control policy should be taken into account when considering any changes to physician compensation.

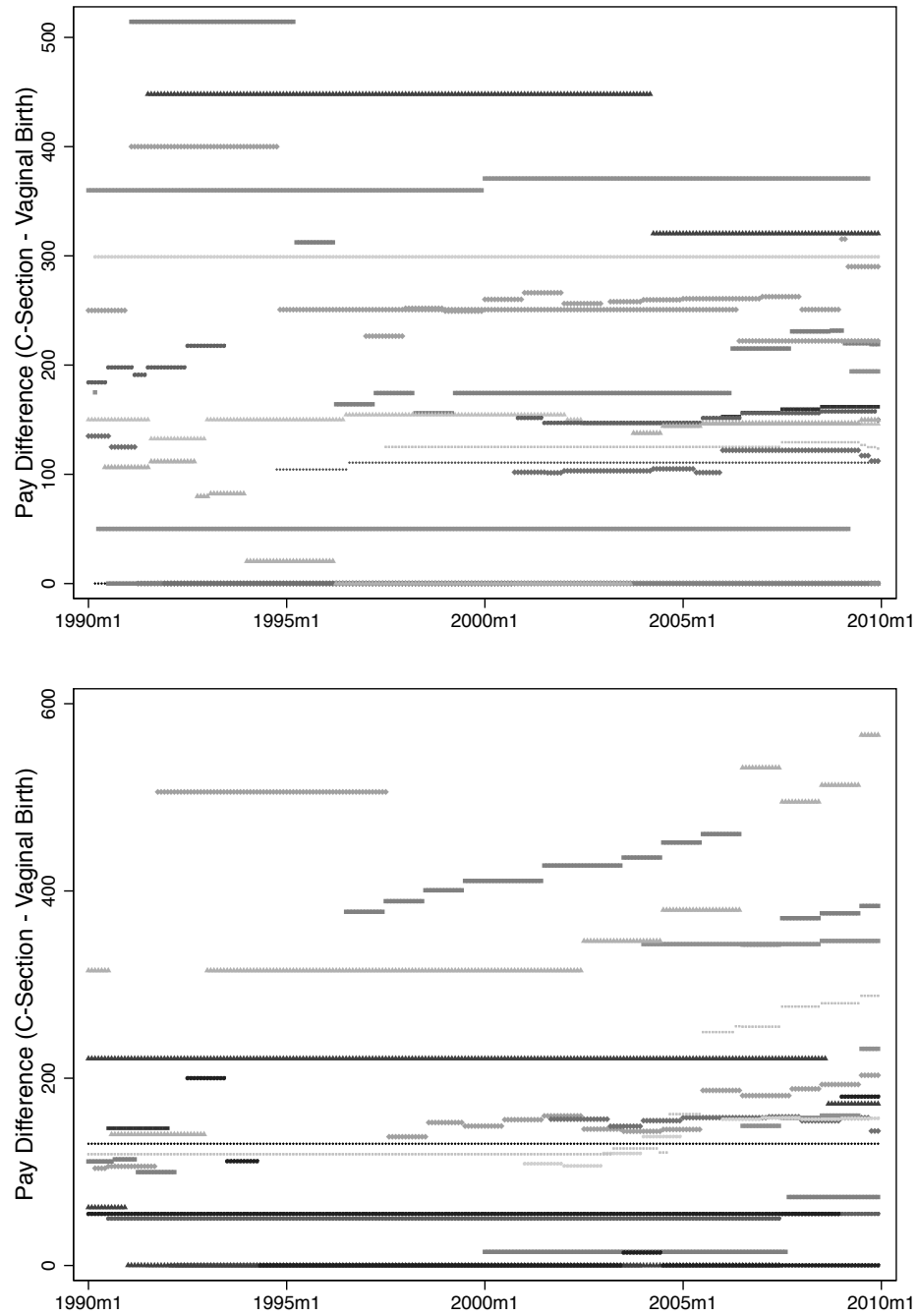
3.8 Figures

Figure 3.1: Medicaid Coverage of Births in the US



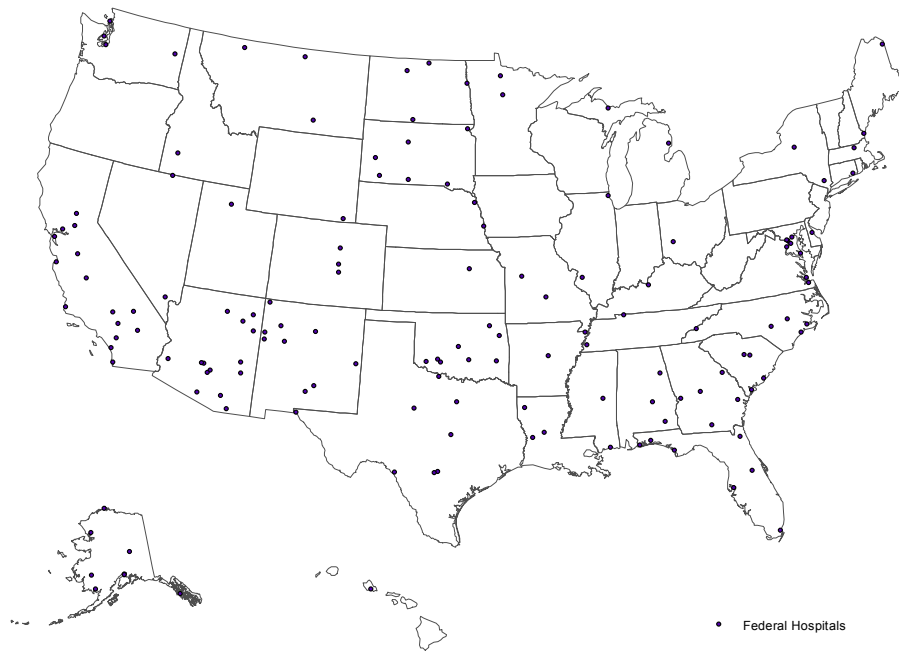
Notes: Data from the Healthcare Cost and Utilization Project.

Figure 3.2: Variation in Pay Difference over Time



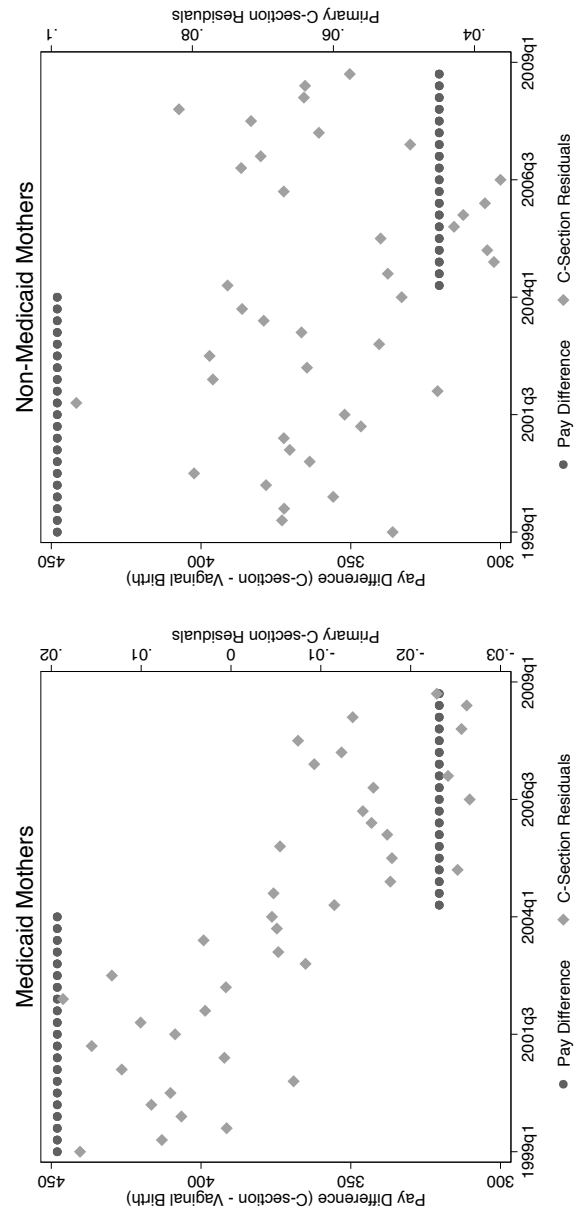
Notes: The top graph displays states Alabama to Mississippi, the bottom graph shows Missouri to Wyoming. Top and bottom one percent trimmed, in order to better display the fee difference variation over time.

Figure 3.3: Locations of Federal Government Controlled Hospitals



Notes: All federally controlled hospitals from 1990-2008.

Figure 3.4: Case Study: Arkansas



Notes: The residuals in Figure 3.4 are from a regression of C-section on year fixed effects for the whole sample, plotted over time for Medicaid and non-Medicaid mothers in Arkansas.

3.9 Tables

Table 3.1: Payment Variable Means

Variables	Mean	SD
Payment for Vaginal Birth	\$788	\$262
Payment for C-Sections	\$922	\$315
Fee Difference (C-Section-Vaginal)	\$135	\$154

Notes: Using CPT codes bundling delivery and postpartum care

Table 3.2: Summary Statistics

Panel A: Individual Medical Risk Factors

Variables	All Mothers	Predicted	Predicted
	Mothers	Medicaid	Private
% Fetal distress	5.5	6.1	5.2
% Smoke during pregnancy	12.5	19.2	4.7
% Lung disease	1.0	1.1	0.9
% Diabetes	3.0	2.4	3.1
% Chronic Hypertension	0.8	0.6	0.8
% Breech	4.1	3.3	5.0
% Low birth weight (< 2501 g)	7.8	9.6	6.5
% Preterm birth (< 37 weeks)	11.8	14.2	9.8
% of Mothers ≤ 16	2.2	5.9	0.0
% of Mothers 40+	2.0	0.6	3.8
Primary C-section rate	17.8	14.9	21.4

Panel B: County Characteristics

Variables	All	High	Low	Counties	Counties
	Counties	Medicaid	Provider	w/o Fed.	w/ Fed.
	Counties	Counties	Counties	Hospitals	Hospitals
% of Mothers ≤ 16	2.3	2.9	2.3	2.4	2.4
% of Mothers 40+	2.0	2.0	2.0	2.0	2.0
% White	79.6	77.8	79.7	80.0	74.9
% Black	14.7	16.1	14.7	14.8	16.0
% Native American	1.1	1.5	1.0	1.0	2.2
% Hispanic	20.1	29.0	20.2	19.9	30.2
% $\geq$ College	22.4	15.3	22.6	22.6	20.8
% Married	63.8	57.8	64.0	63.8	63.9
Median County Income	\$45,060	\$40,893	\$45,139	\$44,857	\$45,456
County Population	105,619	73,220	111,253	102,218	323,372
Avg. unemployment rate	5.2	5.5	5.2	5.3	5.3

Notes: All means in Panel B are population weighted, except for county population.

Table 3.3: Probability of Primary C-section by Insurer in 2012 (x100)

	(1)	(2)	(3)	(4)
Medicaid	-3.0*** (0.1)	-2.4*** (0.1)	-2.3*** (0.1)	-2.8*** (0.1)
Observations	2,573,985	2,573,985	2,573,985	2,573,985
Avg. C-section Rate	21.9	21.9	21.9	21.9
	(1)	(2)	(3)	(4)
Medicaid + Self Pay	-3.4*** (0.1)	-3.1*** (0.1)	-3.0*** (0.1)	-3.5*** (0.1)
Observations	2,698,904	2,698,904	2,698,904	2,698,904
Avg. C-section Rate	21.6	21.6	21.6	21.6
Demographics	X	✓	✓	✓
Medical Risk Factors	X	X	✓	✓
State & Time Fixed Effects	X	X	X	✓

Notes: Omitted category is private insurance. Demographic variables: dummies for Black, American Indian/Alaskan Native, Hispanic, married, education categories, and age in five year bins. Medical risk factors: diabetes, chronic hypertension, eclampsia, breech birth, fetal distress, and an indicator for whether mother smoked during pregnancy (dummies for these variables being missing included when needed, to keep sample constant across specifications). State and time fixed effects: fixed effects for day of week, month of year, year, and state.

Table 3.4: Effect of Salary Structure and Pay Differentials on C-Section Use (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	0.13 (0.18)	0.58*** (0.21)	0.33* (0.17)	0.08 (0.25)	0.50 (0.45)	-0.01 (0.24)
Pay diff. * Salaried	0.02 (0.24)	-0.46** (0.23)	-0.25 (0.21)	0.67 (0.42)	-0.38 (0.39)	0.34 (0.31)
Salaried hospital	-0.28 (0.70)	0.77 (0.69)	0.38 (0.57)	-0.19 (0.64)	-0.04 (1.08)	-0.58 (0.67)
Observations	14,531,737	6,085,151	7,381,948	13,178,385	2,830,188	5,473,547
Mean dept. var	15.0	14.8	14.8	22.0	23.0	21.3

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.5: Effect of Salary Structure and Pay Differentials on Infant Death (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	-0.06*** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)	0.02 (0.01)	0.00 (0.02)	0.01 (0.01)
Pay diff. * Salaried	0.03* (0.02)	0.04** (0.02)	0.05*** (0.02)	-0.01 (0.01)	0.02 (0.02)	0.00 (0.01)
Salaried hospital	0.04 (0.03)	0.03 (0.04)	0.01 (0.03)	0.02 (0.02)	-0.03 (0.03)	-0.01 (0.02)
Observations	14,531,766	6,085,154	7,381,950	13,178,918	2,830,196	5,473,626
Mean dept. var	0.93	0.77	0.90	0.44	0.42	0.47

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.6: Effect of Salary Structure and Pay Differentials on Neonatal Death (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-0.05*** (0.02)	-0.04*** (0.01)	-0.05*** (0.01)	0.01 (0.01)	-0.01 (0.02)	0.01 (0.01)
Pay diff. * Salaried	0.03* (0.01)	0.03** (0.01)	0.04*** (0.01)	-0.00 (0.01)	0.02* (0.01)	0.01 (0.01)
Salaried hospital	0.04 (0.02)	0.05 (0.03)	0.01 (0.02)	0.01 (0.02)	-0.03 (0.03)	-0.01 (0.02)
Observations	14,531,766	6,085,154	7,381,950	13,178,918	2,830,196	5,473,626
Mean dept. var	0.58	0.48	0.55	0.33	0.31	0.35

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.7: Effect of Salary Structure and Pay Differentials on Death After Neonatal Period (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-0.02* (0.01)	-0.02 (0.01)	-0.02** (0.01)	0.01 (0.00)	0.01 (0.01)	0.00 (0.00)
Pay diff. * Salaried	0.01 (0.01)	0.01 (0.01)	0.01* (0.01)	-0.01 (0.00)	-0.01 (0.01)	-0.00 (0.00)
Salaried hospital	0.00 (0.01)	-0.02 (0.02)	-0.00 (0.01)	0.01 (0.01)	-0.00 (0.01)	0.00 (0.01)
Observations	14,531,766	6,085,154	7,381,950	13,178,918	2,830,196	5,473,626
Mean dept. var	0.35	0.29	0.35	0.11	0.11	0.12

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.8: Effect of Pay Difference on Infant Death (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	-0.06** (0.03)	-0.05** (0.02)	-0.06** (0.02)	0.01 (0.01)	0.00 (0.02)	0.00 (0.01)
Pay diff. * High risk	-0.01 (0.02)	-0.05** (0.02)	-0.03* (0.02)	-0.00 (0.01)	0.02 (0.03)	-0.00 (0.02)
High risk	0.14*** (0.03)	0.09** (0.04)	0.10** (0.04)	0.08*** (0.03)	0.01 (0.07)	0.09** (0.04)
Observations	13,172,588	5,368,132	6,345,019	12,083,982	2,360,809	4,769,676
Mean dept. var	0.92	0.76	0.90	0.44	0.41	0.47

Notes: Births in counties with salaried hospitals dropped. All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.9: Effect of Pay Difference on Neonatal Death (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	-0.05** (0.02)	-0.03** (0.01)	-0.04** (0.02)	0.00 (0.01)	-0.01 (0.02)	0.00 (0.01)
Pay diff. * High risk	-0.01 (0.02)	-0.04** (0.02)	-0.03** (0.01)	0.00 (0.01)	0.02 (0.02)	-0.00 (0.01)
High risk	0.13*** (0.03)	0.09** (0.04)	0.08*** (0.03)	0.07*** (0.02)	0.02 (0.06)	0.08** (0.03)
Observations	13,172,588	5,368,132	6,345,019	12,083,982	2,360,809	4,769,676
Mean dept. var	0.58	0.47	0.55	0.33	0.31	0.35

Notes: Births in counties with salaried hospitals dropped. All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.10: Effect of Pay Difference on Death After Neonatal Period (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-0.02*	-0.02	-0.02**	0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Pay diff. * High risk	0.00	-0.01	0.00	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
High risk	0.02	-0.00	0.01	0.01	-0.01	0.01
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Observations	13,172,588	5,368,132	6,345,019	12,083,982	2,360,809	4,769,676
Mean dept. var	0.35	0.29	0.35	0.11	0.11	0.12

Notes: Births in counties with salaried hospitals dropped. All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.11: Effect of Pay Difference on Birth Weight

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-1.95	-8.86***	-3.38	3.06	-5.33	1.90
	(2.80)	(3.03)	(2.40)	(1.93)	(3.69)	(1.83)
Pay diff. * High risk	5.77	7.04	6.62	-3.02	-6.05	-3.47
	(3.90)	(4.69)	(3.83)	(3.49)	(8.76)	(4.04)
High risk	-53.23***	-30.26***	-45.59***	-27.90***	-0.87	-20.00***
	(6.05)	(9.06)	(7.77)	(5.36)	(10.63)	(7.22)
Observations	13,168,875	5,367,052	6,343,434	12,081,362	2,360,577	4,768,957
Mean dept. var	3,214	3,247	3,223	3,364	3,351	3,371

Notes: Births in counties with salaried hospitals dropped. All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.12: Effect of Pay Difference on Gestational Age at Birth

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	0.00	-0.01	-0.01	0.01	-0.01	-0.00
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Pay diff. * High risk	0.01	0.00	0.02	-0.01	-0.03	-0.01
	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.02)
High risk	-0.21***	-0.11***	-0.18***	-0.12***	-0.02	-0.09***
	(2.78)	(3.56)	(3.19)	(2.11)	(3.63)	(2.65)
Observations	12,972,666	5,222,930	6,231,716	12,010,232	2,329,801	4,737,948
Mean dept. var	38.7	38.8	38.7	38.8	38.9	38.8

Notes: Births in counties with salaried hospitals dropped. All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.13: Effect of Pay Difference on Number of Prenatal Visits

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	0.02	0.18**	0.13***	-0.04	-0.04	-0.05
	(0.06)	(0.08)	(0.04)	(0.04)	(0.09)	(0.04)
Pay diff. * High risk	-0.08**	-0.18*	-0.04	0.04	-0.09	0.06*
	(0.04)	(0.11)	(0.04)	(0.03)	(0.09)	(0.03)
High risk	-0.21***	-0.14**	-0.26***	-0.03	0.02	-0.07
	(0.04)	(0.07)	(0.04)	(0.05)	(0.11)	(0.05)
Observations	12,662,987	5,155,846	6,120,307	11,774,693	2,309,217	4,659,290
Mean dept. var	10.4	10.5	10.4	12.3	12.4	12.2

Notes: Births in counties with salaried hospitals dropped. All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

3.10 Appendix

Predicting which mothers are covered by Medicaid at birth I run the following probit regression to predict which mothers are covered by Medicaid in the full sample; the explanatory variables are: married, diabetes, chronic hypertension, black, native american, other race, sex of child, eclampsia, breech, and dummies for state, education categories (four categories and missing), age categories (20-24, 25-29, 30-34, 35-39, 40-44, 45+), hispanic origin (Mexican, Puerto Rican, Cuban, Central American, other and unknown Hispanic, and origin unknown or not stated), birth month, birth order, and week day of birth. In addition, county-year population and population squared, as well as the county-year number of people eligible for Medicaid and the number eligible for Medicaid squared are included, and interacted with race, ethnicity, education, and marital status. The pseudo R^2 is 0.300; prediction based upon 3,388,469 observations from 40 states in the 2012 vital statistics natality file.

Creating an indicator for high risk pregnancies The complications I use for high risk pregnancy varies by time period and state, because of the new birth certificates being phased in, resulting in risk factors being changed, added or dropped. I use all the information I have available for each state and year for the main specification. If I limit the risk factors to only those available throughout the entire time period, my results are largely unchanged.

The total set of complications used to define a high risk pregnancy are: anemia, cardiac disease, acute or chronic lung disease, diabetes, eclampsia, genital herpes, hydramnios/ oligohydramnios, hemoglobinopathy, chronic and pregnancy associated hypertension, incompetent cervix, previous infant weighing over 4000 grams, previous preterm or small-for-gestational-age infant, renal disease, Rh sensitization, uterine bleeding, previous poor pregnancy outcome, other medical risk factors.

Construction of Pay Differential For each pay difference I use in my analysis (C-section minus vaginal birth and C-section minus VBAC), I actually have three different potential measures (see

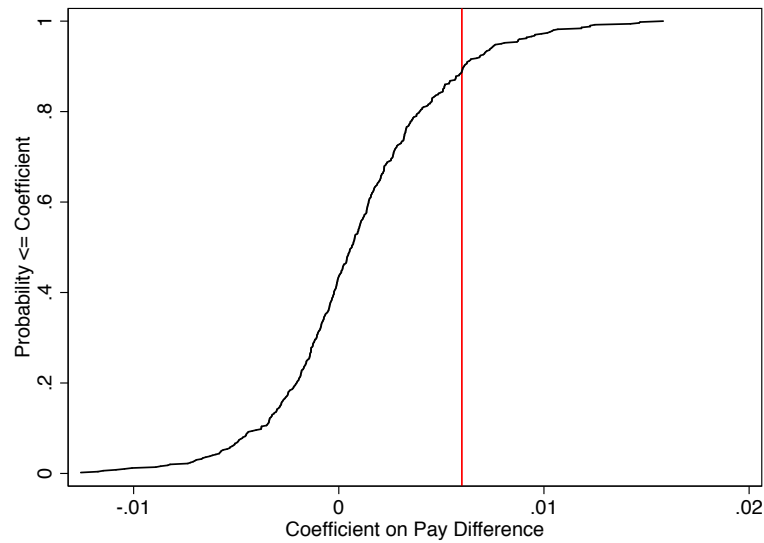
Appendix Table 3.21). Each reflects one of the three ways a delivery procedure can be billed; the bill could be just for the delivery, for the delivery and postpartum care, or for antepartum care, the delivery and postpartum care. I primarily use the pay difference from the codes that include postpartum care, because this variable has the fewest missing state-month observations. In the regressions analysis, I use a composite variable that is constructed as the difference between the “delivery and postpartum care” codes (59515-59410 and 59515-59614), with missing state-month observations filled in using the other two pay difference measures.

Salary Components for Federally Employed Doctors For military doctors, regular compensation consists of officer’s pay based on rank and time in service, basic housing and subsistence allowances, and for those who are eligible, a cost of living allowance. Military physicians do not have to pay for malpractice insurance, and are eligible for special medical bonuses Department of Defense (2012). Doctors in the Indian Health Service are medical officers in the U.S. Public Health Service Commissioned Corps, and face a similar compensation structure. They are also paid a salary, which again increases with promotions and years of service. And like military doctors, IHS doctors also get free clinical practice liability coverage U.S. Public Health Service (2012). The compensation structure of both types of federally employed doctors is characterized by independence of number and type of procedures performed and the inclusion government-paid malpractice insurance.

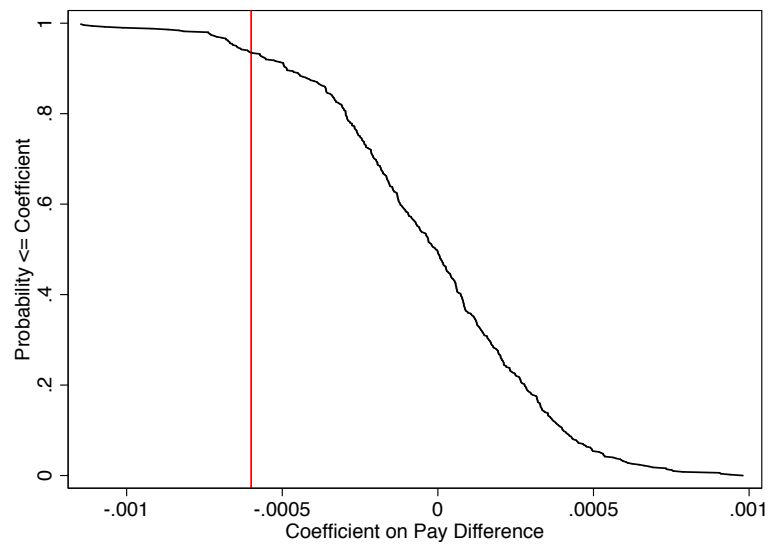
Permutation Analysis States were randomly matched to fee schedule series, and then the main specification was run: Equation 3.2 on the subsample of high Medicaid counties. This was performed 500 times, and the coefficients on the pay difference were recorded. Appendix Figure 3.5 plots the histogram and CDF of the estimates from the permutation analysis; the vertical line shows the estimate reported in the main text.

3.10.1 Appendix Figures

Figure 3.5: Estimates from Randomly Assigning States to Pay Schedules
Estimate for Probability of C-section

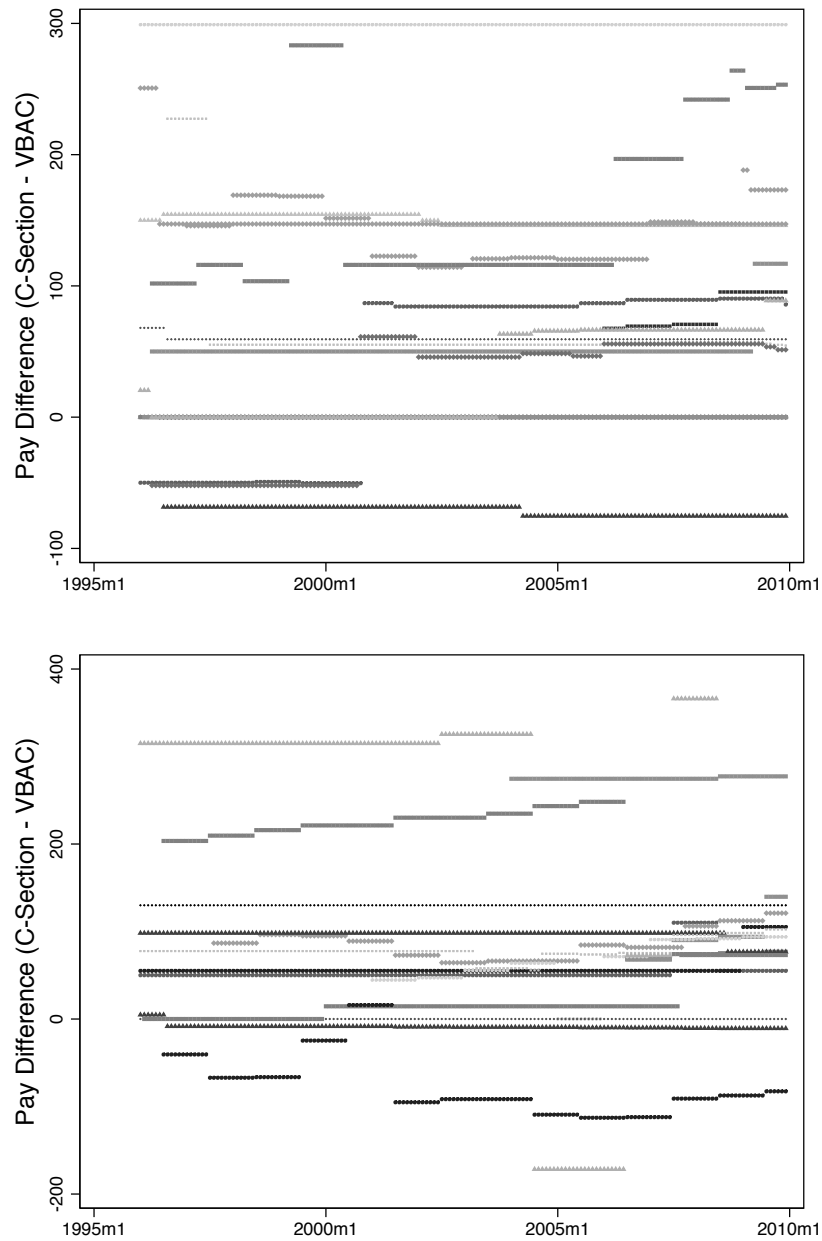


Estimate for Probability of Infant Death



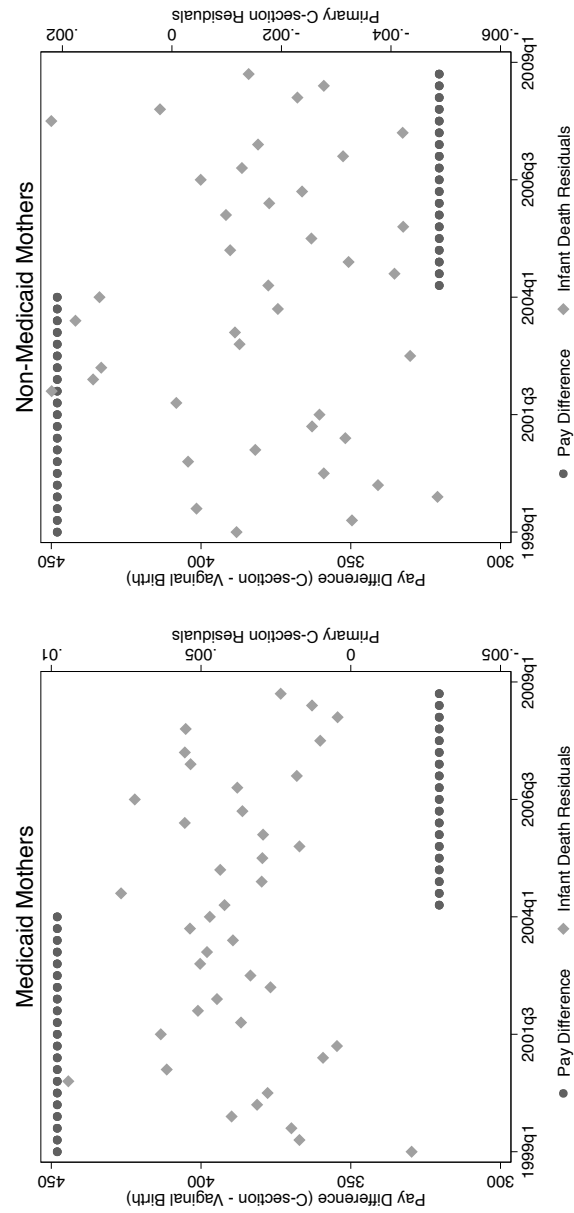
Notes: Each figure is based on 500 repetitions.

Figure 3.6: Variation in Pay Difference over Time



Notes: The top graph displays states Alabama to Mississippi, the bottom graph shows Missouri to Wyoming. Top and bottom one percent trimmed, in order to better display the fee difference variation over time.

Figure 3.7: Case Study: Arkansas (Infant Death)



Notes: The residuals in Figure 3.7 are from a regression of infant death on year fixed effects for the whole sample, plotted over time for Medicaid and non-Medicaid mothers in Arkansas.

3.10.2 Appendix Tables

Table 3.14: Medicaid Fee Schedule Data Used

States	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08
Alabama		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Alaska								x	x	x	x	x	x	x	x	x	x	x	x
Arizona	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Arkansas		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
California	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Colorado								x	x	x	x	x	x	x	x	x	x	x	x
Delaware	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Florida		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Georgia	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Idaho																	x	x	x
Illinois	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Indiana					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Iowa	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Kansas	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Kentucky	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Michigan		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Minnesota	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Mississippi	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Missouri	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Montana	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Nebraska							x	x	x	x	x	x	x	x	x	x	x	x	x
Nevada	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
New Hampshire	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
New Jersey	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
New York	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
North Carolina												x	x	x	x	x	x	x	x
North Dakota																	x	x	x
Oklahoma												x	x	x	x	x	x	x	x
South Carolina	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Texas	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Utah	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Virginia	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Washington			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Wisconsin																x	x	x	x
Wyoming	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Washington, D.C.	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 3.15: Components of Salary for Military and IHS Doctors

Military Doctors	
Variable Special Pay	Monthly bonus received by all physicians including interns and residents
Additional Special Pay	Annual bonus for physicians who graduated a residency or interns serving a GMO tour
Incentive Special Pay	Annual bonus only for physicians who have graduated a residency
Board Certification Pay	Monthly bonus reserved for physicians who are board certified in their medical specialty
Multi-year Special Pay	Annual bonus for doctors who sign a contract for additional years of obligated service
Indian Health Service Doctors	
Retention Special Pay	Annual payment for medical officers who contract to stay on active duty for term of 1+ yrs
Variable Special Pay	Annual special pay based on years of creditable service
Board Certified Pay	Based on years of creditable service and board certification in an accredited specialty
Incentive Special Pay	A special bonus for certain medical officers that is paid annually based on medical specialty
Multi-year Retention Bonus	Payable depending on the specialty training and the duration of the contract

Table 3.16: Women with a Previous C-section: Effect of Salary Structure and Pay Differentials on C-Section Use (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-0.14 (0.27)	-0.04 (0.31)	-0.15 (0.31)	-0.15 (0.19)	0.23 (0.28)	0.08 (0.18)
Pay diff. * Salaried hosp.	-0.43 (0.48)	-0.60 (0.49)	-0.53 (0.54)	0.27 (0.28)	-0.47 (0.31)	-0.11 (0.22)
Salaried hospital	0.28 (1.36)	-0.86 (2.80)	0.97 (1.42)	0.67 (0.72)	1.71 (1.13)	0.19 (0.61)
Observations	1656280	707961	819043	1492034	303556	614811
Mean dept. var	85.6	87.5	86.7	85.6	87.8	86.2

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level. Only births from 1996-2008 used, as VBAC payment code did not exist before this period.

Table 3.17: Women with a Previous C-section: Effect of Pay Differentials on Infant Death (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-0.00 (0.01)	0.00 (0.02)	-0.00 (0.02)	0.01 (0.01)	0.01 (0.02)	0.03 (0.02)
Pay diff. * High risk	-0.03 (0.02)	-0.04 (0.02)	-0.02 (0.03)	-0.00 (0.01)	-0.01 (0.03)	-0.02 (0.02)
High risk	0.00 (0.04)	0.00 (0.05)	0.02 (0.05)	-0.02 (0.03)	-0.18** (0.08)	-0.00 (0.05)
Observations	1522813	639691	713193	1376856	255938	541781
Mean dept. var	0.75	0.64	0.76	0.32	0.31	0.36

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level. Only births from 1996-2008 used, as VBAC payment code did not exist before this period.

Table 3.18: Stratifying by Medicaid Managed Care: Effect of Salary Structure and Pay Differentials on C-Section Use (x100)

Low Managed Care States						
	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	0.32 (0.24)	0.76* (0.46)	0.29 (0.28)	0.18 (0.22)	1.25* (0.69)	0.09 (0.29)
Pay diff. * Salaried hosp.	-0.50 (0.45)	-0.87 (1.26)	-0.07 (0.41)	0.61 (0.44)	0.60 (0.58)	0.78* (0.47)
Salaried hospital	0.41 (1.12)	1.26 (2.52)	-0.63 (0.99)	-1.25 (0.85)	-0.96 (0.98)	-1.71** (0.87)
Observations	3796256	1005058	1664929	3261733	528904	1464714
Mean dept. var	0.1536	0.1608	0.1596	0.2097	0.2185	0.2029
High Managed Care States						
	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	0.07 (0.21)	0.36 (0.25)	0.20 (0.21)	-0.01 (0.33)	-0.63 (0.49)	-0.42 (0.30)
Pay diff. * Salaried hosp.	0.06 (0.22)	-0.18 (0.24)	-0.10 (0.22)	0.17 (0.27)	0.59 (0.41)	0.50** (0.23)
Salaried hospital	-0.16 (0.54)	-0.33 (0.59)	-0.16 (0.56)	-0.76 (0.56)	-2.22* (1.16)	-1.92*** (0.51)
Observations	5280621	3337657	3071404	4999452	1461146	2069471
Mean dept. var	0.1380	0.1384	0.1317	0.2107	0.2218	0.2023

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.19: Effect of Salary Structure and Pay Differentials on C-Section Use: Linear State Trends (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High Medicaid	Low Provider	All	High Medicaid	Low Provider
	Counties	Counties	Counties	Counties	Counties	Counties
Pay difference	0.16 (0.24)	0.40* (0.23)	0.37 (0.31)	-0.11 (0.31)	0.15 (0.56)	0.01 (0.28)
Pay diff. * Salaried	-0.10 (0.27)	-0.46* (0.24)	-0.24 (0.26)	0.80* (0.43)	0.03 (0.51)	0.48 (0.31)
Salaried hospital	-0.10 (0.65)	0.77 (0.80)	0.30 (0.54)	-0.30 (0.48)	-0.52 (1.61)	-0.76 (0.55)
Observations	14,531,737	6,085,151	7,381,948	13,178,385	2,830,188	5,473,547
Mean dept. var	15.0	14.8	14.8	22.0	23.0	21.3

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.20: Effect of Salary Structure and Pay Differentials on Infant Death: Linear State Trends (x100)

	High Predicted Medicaid Mothers			Low Predicted Medicaid Mothers		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Counties	High Medicaid Counties	Low Provider Counties	All Counties	High Medicaid Counties	Low Provider Counties
Pay difference	-0.03 (0.02)	-0.06*** (0.02)	-0.05*** (0.02)	0.02*** (0.01)	-0.01 (0.02)	0.02 (0.01)
Pay diff. * Salaried	0.01 (0.02)	0.03 (0.02)	0.04** (0.01)	-0.02 (0.01)	0.01 (0.02)	-0.01 (0.01)
Salaried hospital	0.05 (0.03)	0.06 (0.05)	0.02 (0.03)	0.03* (0.02)	-0.01 (0.03)	0.00 (0.02)
Observations	14,531,766	6,085,154	7,381,950	13,178,918	2,830,196	5,473,626
Mean dept. var	0.93	0.77	0.90	0.44	0.42	0.47

Notes: All regressions include maternal controls: indicators for race and ethnicity (black, Hispanic, and American Indian/ Alaskan Native), four education categories, age in five year bins, maternal diabetes, eclampsia, breech presentation, and chronic hypertension. In addition, deciles of county-level median income, month and day of week of birth, sex of infant, birth order, and state and year fixed effects are also included. Standard errors clustered at the county level.

Table 3.21: Procedure and Billing Descriptions

Birth Procedures	CPT Code Included in Billing	
Vaginal delivery	59400	Delivery, ante and postpartum care
(w/ or w/o episiotomy	59409	Delivery only
and/or forceps)	59410	Delivery and postpartum care
Cesarean Delivery	59510	Delivery, ante and postpartum care
	59514	Delivery only
	59515	Delivery and postpartum care
VBAC	59610	Delivery, ante and postpartum care
(w/ or w/o episiotomy	59612	Delivery only
and/or forceps)	59614	Delivery and postpartum care
Cesarean delivery, following	59618	Delivery, ante and postpartum care
attempted VBAC	59620	Delivery only
	59622	Delivery and postpartum care

Notes: An episiotomy is a cut made at the opening of the vagina during childbirth, to aid delivery and prevent tissue rupture; forceps are a surgical instrument used to assist the delivery.

Table 3.22: Relationship between Pay Differentials and Unemployment

	(1)
	Pay Difference
	b/se
State unemployment	6.87 (6.17)
L1 State unemployment	0.72 (9.98)
L2 State unemployment	-9.25 (10.08)
L3 State unemployment	1.35 (10.24)
L4 State unemployment	5.20 (6.95)
L12 State unemployment	-0.95 (1.53)
Observations	5709
Mean dept. var	140
Mean unemployment	5.8
State & Year Fixed Effects	✓

Notes: Regression run at the state-month level. L1 denotes state unemployment lagged one month (L2, L3, L4, and L12 defined analogously).

Model This model is a simple extension of the framework from Gruber et al. (1999), where doctors choose between two procedures: C-section and vaginal birth. In this model, the choice variable for doctors is C-section inducement per birth, i . Since inducement has no natural units, I follow Gruber et al. (1999) and define the proportion of total deliveries that are Cesarean, $a(i)$, as linear in i . In this model, $a(0)$ is the C-section rate if there was no inducement.²³ Physicians get utility from income (Y), but disutility from total inducement (I) — the total number of extra C-sections performed. Doctors incur a psychic cost from advising unnecessary surgery, but are reimbursed more for C-sections than for vaginal births, $p_C > p_V$. I assume throughout that $p_C - c_C > p_V - c_V$. This is in line with the fact that at most hospitals, doctors with admitting privileges are not charged by the hospital for resources used, in which case $c_C = c_V = 0$.

To keep things simple, I follow McGuire and Pauly (1991) and Gruber and Owings (1996) and assume that the utility function of physicians is additively separable, with the form

$$U(Y, I) = V(Y) + W(I) \quad (3.5)$$

where Y is income and I is total inducement. The standard assumptions on the utility function hold: $U_Y > 0$, $U_I < 0$, and $U_{YY}, U_{II} < 0$.

Fee-for-Service Compensation Regime The income of a fee-for-service doctor is equal to

$$\begin{aligned} Y &= N (a(i) (p_C - c_C) + (1 - a(i)) (p_V - c_V)) \\ &= Na(i) m + NY_V \end{aligned} \quad (3.6)$$

and I , total inducement, is

²³The C-section rate in the absence of inducement, $a(0)$, is the C-section rate when reimbursements are equalized between procedures. This is not necessarily the optimal rate, as it includes the impact of costs and benefits to doctors that are not reflected in pay— for example, scheduling convenience.

$$I = Ni \quad (3.7)$$

where N is the total mass of mothers giving birth, and i is inducement of C-sections per patient.²⁴ The difference in revenue from performing a C-section versus a vaginal birth is $m = Y_C - Y_V = (p_C - c_C) - (p_V - c_V)$, and $a(i)$ is the share of deliveries that are Cesarean as a function of i . As in Gruber et al. (1999), $a'(i) > 0$, and $a'' = 0$.

For a physician choosing i to maximize utility, the first order condition gives

$$\begin{aligned} U_Y a'(i^*) m + U_I &= 0 \\ a'(i^*) m &= -\frac{U_I}{U_Y} \end{aligned} \quad (3.8)$$

where i^* is the optimally chosen level of inducement per birth. The physician picks i^* so that the marginal return (in \$) of inducing C-sections equals the psychic marginal utility cost of higher inducement. In order to see how procedure choice responds to the pay difference, I differentiate the first order condition fully with respect to m ;

$$\begin{aligned} \frac{di}{dm} &= \frac{-U_{YY} a(i^*) a'(i^*) m - \frac{U_Y a'(i^*)}{N}}{U_{YY} (a'(i^*) m)^2 + U_{II}} \\ &= \underbrace{\frac{-U_{YY} a(i^*) a'(i^*) m}{U_{YY} (a'(i^*) m)^2 + U_{II}}}_{< 0} + \underbrace{\frac{-\frac{U_Y a'(i^*)}{N}}{U_{YY} (a'(i^*) m)^2 + U_{II}}}_{> 0} \\ &\quad \text{income effect} \qquad \text{substitution effect} \end{aligned} \quad (3.9)$$

²⁴It is assumed that the total number of births is fixed. This could be problematic in the context of Medicaid beneficiaries, who doctors may choose not to see. Since I am looking at births, however, I think this assumption is reasonable— if a woman shows up to a hospital in labor, by law she cannot be turned away. Under the Emergency Medical Treatment and Active Labor Act of 1986, hospitals are required to deliver the baby of a woman in active labor, unless the institution is not equipped, as in the case of lacking a neonatal ICU for a high-risk pregnancy.

While the overall sign of $\frac{di}{dm}$ is ambiguous, it separates into the income and substitution effects of a change in the pay difference (m). If $m = Y_C - Y_V$ increases and the substitution effect dominates, then an increase in the pay difference will increase inducement, as the doctor substitutes patients into delivery via C-section. Conversely, if the income effect dominates, the fact that doctors are more wealthy due to the increased pay difference will lead to a decrease in C-sections; inducement will go down.

Salaried Compensation Regime Now consider the case of a salaried doctor. In this case, a physician's income is not impacted by the mix of procedures that is chosen. Thus,

$$Y = \text{Salary} \quad (3.10)$$

and therefore

$$I = 0 \quad (3.11)$$

because the doctor has no financial incentive to induce C-sections, and inducement enters negatively into the utility function. Since Y does not depend on $a(i)$, the doctor's choice of procedure is not influenced by income considerations, and the share of mothers who receive C-sections is $a(0)$.

Bibliography

Abaluck, Jason, Leila Agha, Christopher Kabrhel, Ali Raja, and Arjun Venkatesh, “Negative Tests and the Efficiency of Medical Care: What Determines Heterogeneity in Imaging Behavior?,” Technical Report, NBER Working Paper No. 19956 2015.

Agency for Healthcare Research and Quality, “Hospital Gain-Sharing Program Offers Incentives to Physicians Based on Their Efficiency, Producing Significant Cost Savings Without Decline in Quality,” Technical Report 2014.

Allcott, Hunt, “Site Selection Bias in Program Evaluation,” *The Quarterly Journal of Economics*, 2015, *130*(3), 1117–1165.

Baicker, Katherine, Kasey Buckles, and Amitabh Chandra, “Geographic Variation in the Appropriate Use of Cesarean Delivery,” *Health Affairs*, 2006, *25*(5), 355–367.

Barber, Emma, Lisbet Lundsberg, Kathleen Belanger, Christian Pettker, Edmund Funai, and Jessica Illuzzi, “Indications Contributing to the Increasing Cesarean Delivery Rates,” *Obstetrics and Gynecology*, 2011, *118*, 29–38.

Bishop, Tara, Matthew Press, Salomeh Keyhani, and Harold Alan Pincus, “Acceptance of Insurance by Psychiatrists and the Implications for Access to Mental Health Care,” *JAMA Psychiatry*, 2014, *71* (2), 176–181.

- Björkenstam, Charlotte, Lars-Age Johansson, Peter Nordström, Ingemar Thiblin, Anna Fugelstad, Johan Hallqvist, and Rickard Ljung**, “Suicide or Undetermined Intent? A Register Based Study of Signs of Misclassification,” *Population Health Metrics*, 2014, 12(11), 1–11.
- Black, Sandra, Paul Devereux, and Kjell Salvanes**, “From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes,” *The Quarterly Journal of Economics*, 2007, 122(1), 409–439.
- Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston**, “Will Risk-Adjustment Decrease Health Care Costs? New Evidence from the Medicare Advantage Programs,” Technical Report 2011.
- Buerhaus, Peter, Catherine DesRoches, Robert Dittus, and Karen Donelan**, “Practice Characteristics of Primary Care Nurse Practitioners and Physicians,” *Nursing Outlook*, 2015, 63 (2), 144–153.
- Bureau of Health Workforce**, “Designated Health Professional Shortage Areas Statistics,” Technical Report, US Department of Health and Human Services 2016.
- Cameron, Colin and Douglas Miller**, “A Practitioner’s Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 2015, 50(2), 317–372.
- Case, Anne and Angus Deaton**, “Rising Morbidity and Mortality in Midlife Among White Non-Hispanic Americans in the 21st Century,” *Proceedings of the National Academy of Sciences of the United States of America*, 2015, 112(49), 15078–15083.
- Charlson, Mary, Robert Charlson, Janey Peterson, Spyridon Marinopoulos, William Briggs, and James Hollenberg**, “The Charlson Comorbidity Index is Adapted to Predict Costs of Chronic Disease in Primary Care Patients,” *Journal of Clinical Epidemiology*, 2008, 61(12), 1234–1240.

- Clemens, Jeffrey and Joshua Gottlieb**, “Do Physicians’ Financial Incentives Affect Medical Treatment and Patient Health?,” *The American Economic Review*, 2014, 104(4), 1320–1349.
- Coey, Dominic**, “Physician Incentives and Treatment Choices in Heart Attack Management,” Technical Report, SIEPR Discussion Paper No. 12-027 2013.
- Cullen, Julie Berry, Steven Levitt, Erin Robertson, and Sally Sadoff**, “What Can Be Done to Improve Struggling High Schools?,” *The Journal of Economic Perspectives*, 2013, 27(2), 133–152.
- Currie, Janet and Mark Stabile**, “Child Mental Health and Human Capital Accumulation: The Case of ADHD,” *Journal of Health Economics*, 2006, 25 (6), 1094–1118.
- **and W. Bentley MacLeod**, “First Do No Harm? Tort Reform and Birth Outcomes,” *The Quarterly Journal of Economics*, 2008, 123(2), 795–830.
- **and —**, “Diagnosis and Unnecessary Procedure Use: Evidence from C-Section,” Technical Report, NBER Working Paper No. 18977 2013.
- **, Jonathan Gruber, and Michael Fischer**, “Physician Payments and Infant Mortality: Evidence from Medicaid Fee Policy,” *The American Economic Review*, 1995, 85(2), 106–111.
- Cutler, David**, “Empirical Evidence on Hospital Delivery under Prospective Payments,” Technical Report, MIT 1990.
- **,** “The Incidence of Adverse Medical Outcomes under Prospective Payment,” *Econometrica*, 1995, 63(1), 29–50.
- Dafny, Leemore**, “How Do Hospitals Respond to Price Changes?,” *American Economic Review*, 2005, 95(5), 1525–1547.
- Dehejia, Rajeev and Adriana Lleras-Muney**, “Booms, Busts, and Babies’ Health,” *The Quarterly Journal of Economics*, 2004, 119(3), 1091–1130.

Department of Defense, “Special and Incentive Pay Index,” 2012.

Department of Health and Human Services, “Better Care. Smarter Spending. Healthier People: Paying Providers for Value, Not Volume,” Technical Report 2015.

Dickstein, Michael, “Physician vs. Patient Incentives in Prescription Drug Choice,” Technical Report 2014.

Dranove, David and Paul Wehner, “Physician-Induced Demand for Childbirths,” *The Journal of Health Economics*, 1994, 13(1), 61–73.

—, **Daniel Kessler, Mark McClellan, and Mark Satterthwaite**, “Is More Information Better? The Effects of “Report Cards” on Health Care Providers,” *Journal of Political Economy*, 2003, 111(3), 555–588.

Duflo, Esther, “Scaling up and Evaluation,” Technical Report, Annual World Bank Conference on Development Economics 2004.

Duggan, Mark, “Does Contracting Out Increase the Efficiency of Government Programs? Evidence from Medicaid HMOs,” *Journal of Public Economics*, 2004, 88(12), 2549–2572.

— **and Tamara Hayford**, “Has the Shift to Managed Care Reduced Medicaid Expenditures? Evidence from State and Local-Level Mandates,” *Journal of Policy Analysis and Management*, 2013, 32(3), 505–535.

Ellis, Randall and Thomas McGuire, “Hospital Response to Prospective Payment: Moral Hazard, Selection, and Practice-Style Effects,” *The Journal of Health Economics*, 1996, 15, 257–277.

Ettner, Susan, Richard Frank, and Ronald Kessler, “The Impact of Psychiatric Disorders on Labor Market Outcomes,” *Industrial and Labor Relations Review*, 1997, 51 (1), 64–81.

Families USA, “Presumptive Eligibility: A Step Toward Streamlined Enrollment in Medicaid and CHIP,” Technical Report, Issue Brief 2011.

- Foo, Patricia, Robin Lee, and Kyna Fon**, “Physician Prices, Hospital Prices and Treatment Choice in Labor and Delivery,” Technical Report 2015.
- Grant, Darren**, “Physician Financial Incentives and Cesarean Delivery: New Conclusions from the Healthcare Cost and Utilization Project,” *Journal of Health Economics*, 2009, 28(1), 244–250.
- Greenberg, Greg and Robert Rosenheck**, “Jail Incarceration, Homelessness, and Mental Health: A National Study,” *Psychiatric Services*, 2008, 59 (2), 170–177.
- Gruber, Jonathan**, “Covering the Uninsured in the United States,” *Journal of Economic Literature*, 2008, 46(3), 51–68.
- , “The Cost Implications of Health Care Reform,” *The New England Journal of Medicine*, 2010, 362, 2050–2051.
- **and Maria Owings**, “Physician Financial Incentives and Cesarean Section Delivery,” *RAND Journal of Economics*, 1996, 27(1), 99–123.
- , **John Kim, and Dina Mayzlin**, “Physician Fees and Procedure Intensity: the Case of Cesarean Delivery,” *Journal of Health Economics*, 1999, 18(4), 473–490.
- Hadley, Jack, James Reschovsky, Catherine Corey, and Stephen Zuckerman**, “Medicare Fees and the Volume of Physicians’ Services,” *Inquiry*, 2009, 46(4), 372–390.
- , **Jean Mitchell, and Jeanne Mandelblatt**, “Medicare Fees and Small Area Variations in Breast-Conserving Surgery among Elderly Women,” *Medical Care Research and Review*, 2001, 58(3), 334–360.
- Hartley, David, Valerie Hart, Nancy Hanrahan, and Stephenie Loux**, “Are Advanced Practice Psychiatric Nurses a Solution to Rural Mental Health Workforce Shortage?,” Technical Report, Maine Rural Health Research Center, Working Paper 31 2004.

- Healthcare Cost and Utilization Project**, “Statistical Brief 110: Hospitalizations Related to Childbirth, 2008,” Technical Report, Agency for Healthcare Research and Quality 2012.
- Heberlein, Martha, Tricia Brooks, Jocelyn Guyer, Samatha Artiga, and Jessica Stephens**, “Holding Steady, Looking Ahead: Annual Findings of a 50-State Survey of Eligibility Rules, Enrollment and Renewal Procedures and Cost-Sharing Practices in Medicaid and CHIP, 2010-2011,” Technical Report, Kaiser Commission on Medicaid and the Uninsured 2011.
- Hilkevitch, Jon**, “Suicide is Top Cause of Train Track Deaths,” *Chicago Tribune*, March 3rd, 2005.
- Hillman, Bruce and Jeff Goldsmith**, “The Uncritical Use of High-Tech Medical Imaging,” *The New England Journal of Medicine*, 2010, 363, 4–6.
- Ho, Kate and Ariel Pakes**, “Hospital Choices, Hospital Prices and Financial Incentives to Physicians,” Technical Report, NBER Working Paper No. 19333 2014.
- Johnson, Erin and M. Marit Rehavi**, “Physicians Treating Physicians: Information and Incentives in Childbirth,” Technical Report 2013.
- Keeler, Emmett and Thomas Fok**, “Equalizing Physician Fees Had Little Effect on Cesarean Rates,” *Medicare Care Research and Review*, 1996, 53, 465–471.
- Kolstad, Jonathan**, “Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards,” *American Economic Review*, 2013, 103(7), 2875–2910.
- Lee, Henry, Jeffrey Gould, W. John Boscardin, Yasser El-Sayed, and Yair Blumenfeld**, “Trends in Cesarean Delivery for Twin Births in the United States: 1995-2008,” *Obstetrics and Gynecology*, 2011, 118(5), 1095–1101.
- Leibowitz, Arleen, Joan Buchanan, and Joyce Mann**, “A Randomized Trial to Evaluate the Effectiveness of a Medicaid HMO,” *Journal of Health Economics*, 1992, 11(3), 235–257.

- Lessler, Adam, Joshua Isserman, Rajan Agarwal, Harold Palevsky, and Jesse Pines**, “Testing Low-Risk Patients for Suspected Pulmonary Embolism: A Decision Analysis,” *Annals of Emergency Medicine*, 2010, 55(4), 316–326.
- MacDorman, Marian, Eugene Declercq, and Fay Menacker**, “Recent Trends and Patterns in Cesarean and Vaginal Birth After Cesarean (VBAC) Deliveries in the United States,” *Clinics in Perinatology*, 2011, 38(2), 179–192.
- Manning, Willard, Joseph Newhouse, Naihua Duan, Emmett Keeler, and Arleen Leibowitz**, “Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment,” *American Economic Review*, 1987, 77(3), 251–277.
- McGuire, Thomas and Mark Pauly**, “Physician Response to Fee Changes with Multiple Payers,” *Journal of Health Economics*, 1991, 10, 385–410.
- Milner, A. D., R. A. Saunders, and I. E. Hopkin**, “Effects of Delivery by Caesarean Section on Lung Mechanics and Lung Volume in the Human Neonate,” *Archives of Disease in Childhood*, 1978, 53, 545–548.
- Office of Inspector General**, “Special Advisory Bulletin: Gainsharing Arrangements and CMPs for Hospital Payments to Physicians to Reduce or Limit Services to Beneficiaries,” Technical Report 1999.
- Orszag, Peter and Ezekiel Emanuel**, “Health Care Reform and Cost Control,” *New England Journal of Medicine*, 2010, 363(7), 601–603.
- Quan, Hude, Bing Li, and Chantal Couris**, “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*, 2011, 173(6), 676–682.
- Rao, Vijay and David Levin**, “The Overuse of Diagnostic Imaging and the Choosing Wisely Initiative,” *Annals of Internal Medicine*, 2012, 157 (8), 574–576.

- Reinhardt, Uwe**, “The Pricing of U.S. Hospital Services: Chaos Behind a Veil of Secrecy,” *Health Affairs*, 2006, 25, 57–69.
- Rockett, Ian, Julie Samora, and Jeffrey Coben**, “The Black-White Suicide Paradox: Possible Effects of Misclassification,” *Social Science and Medicine*, 2006, 63, 2165–2175.
- Ruhm, Christopher**, “Are Recessions Good for Your Health?,” *The Quarterly Journal of Economics*, 2000, 115(2), 617–650.
- Sekunda, Joanne Rawlings, Deborah Curtis, and Neva Kaye**, “Emerging Practices in Medicaid Primary Care Case Management Programs,” Technical Report, National Academy for State Health Policy 2001.
- Silverman, Elaine and Jonathan Skinner**, “Medicare Upcoding and Hospital Ownership,” *Journal of Health Economics*, 2004, 23, 369–389.
- Smith-Bindman, Rebecca**, “Is Computed Tomography Safe?,” *New England Journal of Medicine*, 2010, 363 (1), 1–4. PMID: 20573919.
- Smith, Sherry, ed.**, *Medicare RBRVS 2012: The Physician Guide*, American Medical Association, 2012.
- Substance Abuse and Mental Health Services Administration**, “Treatments for Mental Disorders,” Technical Report 2015.
- Timbie, Justin, Peter Hussey, Lane Burgette, Neil Wenger, Afshin Rastegar, Ian Brantley, Dmitry Khodyakov, Kristin Leuschner, Beverly Weidmer, and Katherine Kahn**, “Medicare Imaging Demonstration Final Evaluation: Report to Congress,” Technical Report, RAND Corporation 2014.
- U.S. Public Health Service**, “Salary and Benefits,” January 2012.

Yip, Winnie, “Physician Responses to Medical Fee Reductions: Changes in the Volume and Intensity of Supply of Coronary Artery Bypass Graft (CABG) Surgeries in the Medicare and Private Sectors,” *Journal of Health Economics*, 1998, 17, 675–699.