

Essays on the Evolution of Healthcare Technology

by

Ashley Hodgson

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Alan Auerbach, Chair
Professor Emmanuel Saez
Associate Professor Richard Scheffler

Spring 2011

Essays on the Evolution of Healthcare Technology

Copyright 2011
by
Ashley Hodgson

Abstract

Essays on the Evolution of Healthcare Technology

by

Ashley Hodgson

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Alan Auerbach, Chair

This dissertation looks at health care technology using the tools and methods of economics. The particular focus is on the causes and implications of dynamic changes in health care technology over time. The dissertation utilizes methodologies developed in public finance to consider the effects of policy change on technological innovation and adoption.

I first look at the impact of Medicare's prospective payment system and how it influences which technologies get developed and adopted. This presents an ideal case study because the government implemented the system nationwide in 1983, and there have been many years for innovators to respond to the different financial incentives. Prospective payment theoretically penalizes hospitals for adopting technologies that treat illnesses common among the elderly. This chapter evaluates whether we see empirical evidence that there has been fewer innovative developments targeting illnesses common among the elderly compared to illnesses common among the non-elderly. The data paint a picture that supports the theoretical predictions, and upholds the idea that payment incentives do indeed impact which technologies get developed in the first place.

The next chapter looks at a much smaller government change and its short run effect on hospital behavior. Every year, the government adds a few new procedures to the list of icd-9-cm codes. These codes make it easier for hospitals to bill insurers for procedures. This chapter investigates empirically and finds that there is a sudden jump in the number of procedures performed in the quarter when a new code is introduced and that this jump persists going forward. It also looks at different sub-groups of insurers and hospitals, and finds that patients whose insurers depend most heavily on the icd-9-cm codes have the largest jump in the probability of undergoing the procedure in the quarter when it is introduced. The jump in treatment is non-existent for Medicare patients and self-pay patients, for whom the icd-9-cm procedure code is irrelevant.

The final chapter investigates changes in ADHD medication over time. Understanding these changes is important in understanding the diffusion of new technology. This project looks at a time period, 2001 to 2003, when a long-acting version of ADHD medication

was spreading, which makes it interesting from a technological standpoint. The project particularly asks why some counties have higher growth rates in medication than other counties. What factors lead to faster diffusion in a particular geographic region? My co-authors and I find that supply side characteristics, such as physicians per capita and a younger age distribution among physicians, leads to a faster rate of diffusion.

These findings together shed light on some specific key questions about health care technology and how it changes over time. These issues will become increasingly important as health care costs escalate and as policymakers strive to make health care more affordable. Economists have long claimed that new technology plays the biggest role in cost growth. While new technology brings benefits as well as costs, economists will need more tools for evaluating technology cost-effectiveness if cost containment becomes an important enough political goal. This project sheds light on some of the matters that will need to be fleshed out in greater detail if we eventually want to understand innovations role in rising costs.

Contents

1	Introduction and Overview	1
2	Does Medicare’s Prospective Payment System Discourage Investment in Therapeutic Technology?	5
2.1	Introduction	5
2.2	Review of Literature	7
2.3	Background on the Relationships between Doctors, Hospitals and Insurers	8
2.4	Theoretical Model	10
2.4.1	Overview of Model	10
2.4.2	Hospital’s maximization problem	10
2.4.3	Solving the Model	15
2.5	Empirical Strategy	16
2.5.1	Data	16
2.5.2	Empirical Model	17
2.6	Results	23
2.7	Robustness checks	34
2.8	Conclusion	40
3	Do Changes in Government Procedure Codes Impact Patient Treatment Patterns?	41
3.1	Introduction	41
3.2	Background and Literature	42
3.3	Methodology	43
3.3.1	Data	43
3.3.2	Empirical Model	44
3.4	Results	47
3.5	Robustness Checks	52
3.6	Conclusion	55

- 4 Health Care Supply and County-Level Variation in ADHD Prescription Medications 57**
- 4.1 Introduction 58
- 4.2 Methods 59
 - 4.2.1 ADHD Prevalence by County 61
 - 4.2.2 Dependent Variable 61
 - 4.2.3 Health Care 63
 - 4.2.4 Sociodemographic Variables 63
 - 4.2.5 School Variables 63
 - 4.2.6 Data Analytic Procedures 64
- 4.3 Results 64
- 4.4 Discussion 65

- Bibliography 70**

List of Figures

2.1	Graphical representation of the hospital's maximization problem	13
2.2	Medicare patients as source of variation	14
2.3	Price discrimination through waiting	15
2.4	Number of Procedures and Diagnoses Per Patient by Age	19
2.5	Procedures Per Patient by Age – Restricted to ages 60 to 70	20
2.6	Average number of therapeutic procedures per patient	24
2.7	Average number of therapeutic procedures per patient	25
2.8	Average number of diagnostic procedures per patient	26
2.9	Average number of therapeutic procedures per patient	27
2.10	Coefficient on number of elderly	31
2.11	Coefficient on number of elderly	32
2.12	Coefficient on number of elderly	33
2.13	Scatter plot of therapeutic procedures v. percent elderly for different categories of illnesses	34
3.1	Probability that a patient undergoes a procedure	46
3.2	Using patients actually diagnosed	54
3.3	Using patients among the top 4% most likely to be diagnosed	54

List of Tables

2.1	Examples of diagnostic and therapeutic procedures	11
2.2	Summary of predictions from the theoretical model	11
2.3	Reason for using hospital fixed effects	21
2.4	Reason for checking analysis without using hospital fixed effects	22
2.5	(2008) Effect of elderly on procedures	28
2.6	(1983) Effect of elderly on procedures	29
2.7	Regressing within each category.	35
2.8	2008 Robustness checks: Different specifications	36
2.9	2008 Robustness checks: Excluding outliers.	37
2.10	1983 Robustness checks: Different specifications	38
2.11	1983 Robustness checks: Excluding outliers	39
3.1	Predicted effect for different sub-groups	45
3.2	Using patients actually diagnosed	48
3.3	Using patients among the top 4% most likely to be diagnosed (predicted) . .	49
3.4	Using patients actually diagnosed	50
3.5	Using patients among the top 4% most likely to be diagnosed (predicted) . .	51
4.1	List of Data Sources	60
4.2	ADHD medication prescriptions	61
4.3	Descriptive county characteristics, 2001 to 2003.	65
4.4	Fixed effects regression results	66

Acknowledgements

I would especially like to thank Alan Auerbach, my adviser. I cannot say enough amazing things about Alan as an adviser. He has allowed me a lot of room for creativity in perusing projects, and at the same time guides the project with rigor in economic methods. That is a really difficult balance to strike, but when an adviser does it as well as Alan does, it preserves both the trust and the fascination in economics that lead us to graduate school in the first place. I am so thankful for Alan's availability, and how often I have been able to meet with him over the years. That interaction was pivotal for me to develop ideas, make mistakes, and find my way out of them. His guidance has strengthened me to stand on my own feet as a researcher. My gratitude for Alan will continue through my career, because it has shaped a large part of who I am as an economist.

Richard Scheffler also played an important role in shaping me as a health economists. He showed me the ropes of the health policy world and got me excited about this subject matter. I am so thankful to have had the opportunity to work with him on so many different levels – on his book, on his research, as a co-author, and on committees. Richard fosters a wonderful sense of community among the people he leads, and that heightens the quality of the research as we feed from one another and sharpen each other's thoughts. I am lucky to have been a part of all that goes on at the Petris Center.

Other Berkeley professors have also helped along the way. Emmanuel Saez has been very honest with me even when it will be hard to hear, and the courage he has shown in doing that has made a truly positive difference in my life. I'm so thankful he steered me away from pure theory into empirical work. Will Dow has taken a few occasions to read my research and think about it very thoroughly, and those sessions with him made a huge difference.

I am thankful to all of my co-authors for the ADHD paper. I learned a lot about Stata and about the process of thinking through an econometric problem from the countless hours I spent with Tim Bruckner. He is an incredible thinker. Brent Fulton is the most detail-oriented person I have ever known, and I have a lot of trust and respect in any research project he is involved in. Chris' persistence in getting all of the data together is admirable, and Peter's practical physician knowledge made the project complete.

I would also like to thank Dan Acland and Anne Moore. I would not have survived the first year without the good luck of getting into a study group with such an amazing teacher as Dan. Both Dan and Anne were part of a dissertation accountability with me one semester

and they kept that group on track, making sure we steered clear of complaining and simply stated our progress and goals over the past week. Their friendship has been valuable on so many levels in the years working on the dissertation. In everything from long conversations at the kitchen table to L^AT_EX questions, they have supported me.

My accountability partner, Sam Fitzgerald, has been so important to me. She prods me to maintain integrity in all areas of my life, including in research. She extends grace when that is what I need to continue forth, and she calls me out when I fall off track. In addition, the various members of my Veritas bible study have been incredible support over the years. I kept through the dissertation process heart because of them.

Finally, my family has supported . Talking with my parents every Sunday has helped me keep life in perspective. Their encouragement means a lot to me. My brother's support is mainly encouragement in writing in general. His own passion for writing is an encouragement to love writing, which has played out in even sometimes enjoying the writing process of the dissertation. And Brad is just an amazing brother in general.

Chapter 1

Introduction and Overview

Health care is not a static product. As a matter of fact, medical innovators labor away today to determine what exactly “health care” will mean 20 years hence. Yet, as health care morphs on a monthly basis, the right to it remains unchanged. It is widely considered an American right, if not a human right. And as the product changes, the cost of this human right compounds. Its’ growing cost could eventually even threaten the right itself. So what can we learn about the changing nature this elusive product, as economists must consider it?

How do dollars get channeled toward one medical cure and away from another? Perhaps scientists and policymakers overlook the remedies that “might have been”. What can we know about an opportunity cost not actuated? Economists are the ones who care and have the tools to ask and answer such questions. So it falls to our lot to study medical innovation. It is an important charge, because if society never looks at what “might have been”, we will never know how to make things better.

Economists recognize that the payment climate and financial potential will determine where the innovators go, and therefore where medical science itself will follow. This, in turn, determines the cost and quality for years to come. Financial incentives as they play out in the evolution of health care will become more and more important to understand as growing costs threaten the rights we have stood by in recent generations.

Policymakers have only just begun to glimpse the magnitude of the problem of growing cost in health care. The new health care bill passed in 2010 focuses primarily on issues of equity in health care, rather than growing costs. Only a few provisions in the law do target cost growth. This includes a tax penalty for high cost plans, limits in the share of health insurance dollars covering overhead, and pricing guidelines in the market for individual insurance. At best, however, these regulations will lead to one-time reductions in cost that will quickly be swamped if cost growth continues at current levels. The bill does address the adverse selection problem, which could also bring a one-time reduction in the average cost of health care by spreading costs across more people.

As people are required to purchase insurance starting in 2014, keeping insurance afford-

able will likely become a burning topic for policymakers going forward. Price controls may not be sufficient or sustainable as a long run solution. In which case, our understanding of the mechanisms behind cost growth in health care will need to expand quickly enough to provide policymakers with answers and, hopefully, tools. The sustainability of Medicare, US health care generally, and potentially even the economy may depend on it. Prior to the time Congress passed the 2010 Health Care Bill, the CBO estimated that 37 percent of GDP would be devoted to health care by 2050. While the CBO has updated government expenditure projections, it has not, to my knowledge, released a new estimate of growth in the health care sector among the larger population. Undoubtedly projection estimates will change once economists factor in anticipated effects of the bill. Still, the prior estimates show the potential danger of the healthcare sector overtaking a large share of the economy.

For the past decade, economists have recognized the importance of new medical technology in shaping the cost growth in health care. Although it is difficult to directly measure the impact of new technology on the rising cost of health care, economists have estimated it indirectly by teasing out all other identifiable contributors. In separate papers, both David Cutler and Joseph Newhouse estimated the contribution of other factors to rising health care costs: an aging population, rising prices, moral hazard, rising incomes, increased administration and supplier induced demand. Cutler estimated about a 50% residual and Newhouse a 70% residual. They attribute this residual to new technology, and economists have long agreed that new technology will lead to price inflation as well.

Of course, rising prices may still be worthwhile if the value of the new technology outweighs the problems brought on by rising costs. Any assessment of new medical technology will need to consider both costs and quality. The study of cost-effectiveness, or effectiveness per dollar, has come onto the stage in the last 20 years as an important tool for assessing new technology. Still, this literature generally takes a static approach rather than a dynamic assessment.

Economists have only recently begun to concentrate efforts at investigating the causes and effects of medical innovation. Therefore, the field remains wide open, with many different avenues of unanswered questions. What policy tools can be used to channel innovation toward the most cost-effective remedies? What constitutes cost-effectiveness in a new technology? What is the ideal rate of change? What is the relationship between diffusion of a new technology and its cost impact on health insurance? What mechanisms propel forward a particular technology's diffusion over competing technologies, old or new? This dissertation makes a few small steps in answering these questions empirically. The focus is on policy tools and technology diffusion.

Diffusion of new technology will be intimately tied in with cost. Following Tobin's q-theory, instant diffusion of new technology would be very costly. The faster something diffuses the more costly the process. In the medical care sector, diffusion costs include the education of doctors, training of new technicians, development of machines and dissemination of information regarding that technology. The conundrum in health care is that once a technology has been discovered, it is perceived to become a basic human right by means of

medical necessity. This places a very high pressure for fast diffusion, which could elevate costs extensively.

What, then, keeps diffusion costs from becoming too high? For one, it may take time for knowledge, understanding and trust of a new technology to spread across doctors. Apart from slow spread of knowledge, however, the main thing keeping costs from escalating is various forms of rationing. This can include rationing by price, waiting lines, delayed payments, refusal of reimbursement and many other innovative forms that managed care organizations have devised in recent years. As information travels more and more quickly via the internet, medical journals, and increased doctor specialization, pressures of increasing diffusion costs could escalate in a problematic way. As such, diffusion and cost growth in health care may be intimately linked. Clearly a study of medical care innovation and rising costs will need to include a look at the causes and consequences of the diffusion of new technologies into the medical field.

This dissertation looks at medical innovation, diffusion of new medical technologies and policy issues surrounding both. Each chapter looks at a particular aspect of medical technology, using data currently available. Chapter 2 has a broader scope, tracing the historical implication of a sweeping policy change that occurred 1983. Chapters 3 and 4 look more narrowly at short-run effects of new technology or policy changes.

Chapter 2 traces the effect of payment incentives on which technologies get developed in the first place. In particular, it looks at the historical implementation of Medicare's prospective payment system in hospitals and tracks the theoretical and empirical implications on medical innovations over the years. It finds support that the prospective payment system is influencing the balance of diagnostic and therapeutic technologies adopted by hospitals. This is particularly relevant when thinking about ways in which policymakers may influence medical innovation going forward. Since Medicare accounts for a very large share of patients, particularly at the hospital level, Medicare's payment schemes may play an important role in determining investment in new technologies, and therefore the shape of health care in years to come.

Chapter 3 looks more narrowly at the effect of a small change in policy: how the addition of a new procedure code changes the way doctors do medicine. Every year an independent committee updates the list of icd-9-cm procedure codes that are widely used among private insurers. This chapter looks at what happened in the 4th quarter of 2003 when the addition of a new code for a "laparoscopic supracervical hysterectomy" made it easier for hospitals to bill for that procedure. The findings show a sudden and persistent uptick in the number of these procedures performed during the quarter when the code was introduced. Moreover, this uptick was not observed for Medicare patients and self-pay patients, whose reimbursement does not depend on icd-9-cm procedure codes. The chapter discusses the methodology and implications of this finding in detail.

Chapter 4 steps outside the hospital realm to look at the introduction of new technology in the pharmaceutical industry. This chapter looks at county-level variation in attention deficit and hyperactivity disorder (ADHD) drugs during 2001 to 2003. This time period is

particularly relevant for studying the diffusion of new technology, because it is the time period when long-acting ADHD medications took the stage. Prior to 2001, most ADHD medication was short-acting, meaning a patient had to take it three times a day for it to be effective. Long acting medication relies on a time-release technology that allows the patient to take a pill only once a day. This study looks at county-level changes in medications over time. It looks at the supply side characteristics (physicians per capita, age distribution of physicians in a county, etc) and the correlation between these and growth in ADHD medications. It finds that having more doctors and younger doctors in a county will lead to a faster growth in ADHD medication use.

Chapter 2

Does Medicare's Prospective Payment System Discourage Investment in Therapeutic Technology?

Abstract

Medicare's prospective payment system reimburses hospitals based on a patient's diagnoses, not treatments. Therefore, treating Medicare patients reduces hospital profit dollar-for-dollar. Medicare patients are protected from under-treatment because doctors make treatment decisions, and Medicare reimburses doctors by procedure and time. The hospitals control which technologies to adopt. Under prospective payment, hospitals will lose money by adopting a technology that treats an illness common among Medicare patients. Do these incentives influence investment across different kinds of medical technologies? This paper investigates empirically by looking at treatment patterns for illnesses common and uncommon among the elderly, controlling for patient age and number of diagnoses. The results show that a 60-year-old today with an illness common among the elderly receives fewer therapeutic treatments than a 60-year-old with an illness uncommon among the elderly. This result disappears when looking at data prior to 1983, the year Medicare implemented prospective payment. In addition, a privately insured patient with an illness common among Medicaid patients also receives fewer procedures. These results suggest that the payment system of the average patient within an illness group will influence innovation for new treatments of that illness.

2.1 Introduction

Over the past four decades, health care costs per capita have grown twice as fast as GDP per capita (CBO, 2008). Economists commonly point the finger at new technology as the

biggest driver behind such rapid cost growth. In other industries, research and development (R&D) drives down costs. Of course, rising costs might still be preferable if they bring about large improvements in health. In an ideal situation, each dollar spent developing these new technologies would support a project with more health-value per dollar than all alternative investment projects.

However, scant evidence supports the notion that current incentives channel R&D dollars optimally. Imagine, for example, that an innovator could devote resources toward two alternative projects: (a) a technology that would reduce the cost of treating cancer with identical effectiveness or (b) a laser technology that treats toe-nail fungus at a cost of \$1,200 per treatment, such as the PinPointe Footlaser (Singer, 2009). Ten percent of the population suffers from a toe-nail fungus that leaves them unable to wear open-toed shoes. Medical need is a zero-one decision, so people with a medical need to remove yellow fungus are entitled access to available treatments under the current regime. Therefore, from the standpoint of an investor in medical development projects, the toe-nail fungus project may well trump any cost-saving technology. Given the compounding nature of innovation, the implications for cost-effectiveness and long-run affordability look grim.

If policymakers want to slow the pace at which health care costs grow, rechanneling R&D dollars toward cost-effective and cost-reducing technologies will be imperative. What tools do policymakers have for influencing medical innovation? What forces do private sector innovators cater to? This paper investigates the impact of hospital payment schemes by zeroing in on some testable predictions about the different R&D incentives created by Medicare's payment system.

Under prospective payment, the government reimburses hospitals based on a patient's diagnoses, regardless of treatments. Each marginal dollar spent treating the patient reduces a hospital's profit by a dollar. Therefore, the hospital should want to minimize the number of procedures Medicare patients undergo. Patients are protected from under-treatment because the doctors make the patient-care decisions, not the hospitals. Doctors are paid based on their time and the procedures they perform. Following this setup, technology adoption can be modeled as a two-stage game. In the first stage, the hospital decides which technologies to adopt, taking into account the expected physician behavior. In the second stage, the doctor decides which technologies to use to treat the patient. The outcome will be fewer therapeutic technologies adopted for illnesses common among the elderly. This will be true even if doctors choose to treat elderly patients more intensively than non-elderly patients, as Card, Dobkin and Maestas (2008) found to be the case.

Based on this intuition, we can make several testable predictions. First, a 60-year-old today with an illness common among the elderly will receive fewer therapeutic treatments than a 60-year-old with an illness common among the young. From a regression standpoint, we expect to observe a negative relationship between the number of procedures a patient undergoes and the number of elderly sharing that patient's primary diagnosis. Second, this negative relationship should disappear prior to 1983, when Medicare implemented prospective payment. The coefficient should decline after 1983 at a rate that depends on the rate

at which new technology displaces old technology. Between 1983 and 2001, the prospective payment system continued to reimburse for capital inputs on a retrospective basis, so we might not expect the relationship to be fully negative until after 2001. Most strikingly, the time pattern on this coefficient fits the story that prospective payment is driving the numbers (Figure 2.10).

Treatment patterns for diagnostic procedures can act as a kind of control group for comparison. Since prospective payment reimburses by diagnoses, technologies that uncover new diagnoses may still be worthwhile for hospitals to adopt for Medicare patients. Therefore, the patterns I observed for therapeutic procedures should not be mimicked in the data on diagnostic procedures. Indeed, the coefficient on diagnostic procedures shows no particular trend over this time period.

Finally, I use the same analysis to test whether a patient with an illness common among MediCal patients receive fewer therapeutic treatments. MediCal is California's version of Medicaid. Nationwide Medicaid reimbursement rates are known to be significantly lower than most private insurers. The Medicaid population will also be representative of the population that receives charity care. The analysis shows that privately insured patients receive both fewer therapeutic treatments and fewer diagnostic treatments if they have an illness common among MediCal patients. This result affirms the model that hospital payment type and reimbursement rate impact the kinds of technologies developed for a particular illness.

Understanding the drivers of medical R&D will be necessary to predict and influence the future growth of technology and costs. President Obama already proposed setting up a government agency to investigate the cost-efficiency of medical technologies, although it was not part of the 2010 health care bill that Congress passed. How could the government use the information from such an agency to actually influence the evolution of medical technology in favor of cost effectiveness? If research establishes a causal relationship between hospital reimbursement schemes and technological development, then Medicare reimbursement can act as such a tool. In addition, if Medicare's prospective payment system is creating distortions in R&D incentives, the government may want to counterbalance those distortions in some way (adjusting NSF grants, adjusting the patent duration by cost-effectiveness, tax incentives, etc).

Section 2.2 reviews the literature. Section 2.3 provides some background on the relationship between doctors, hospitals and insurers. A theoretical model formalizing the intuition described above is laid out in Section 2.4. The data and empirical strategy appear in Section 2.5. Sections 2.6 and 2.7 present the results and robustness checks. Section 2.8 concludes.

2.2 Review of Literature

To my knowledge, this is the first paper looking at R&D incentives created by the prospective payment system (PPS). However, a number of papers have measured effect of PPS on treatment patterns, number of diagnoses, and capital intensity within hospitals. This section

outlines the findings of the PPS literature.

Studies comparing treatment patterns before and after prospective payment show that hospitals reduced costs without significantly reducing quality after the implementation of prospective payment (e.g. Feder, Hadley, and Zuckerman, 1987). For example, the average length of stay dropped without significant negative impacts on outcomes (Caulam and Gaumer, 1991; Gold, et. al. 1993). One-year survival rates upon entering a hospital were the same before and after 1983 (Cutler, 1995). One study found that hospitals reduced the use of some specific medical technologies immediately after the introduction of prospective payment (Sloan, Morrissey and Valvona, 1988).

Evidence also suggests that prospective payment leads to an increase in the number of diagnoses per patient. We might expect that a system that reimburses based on diagnoses would incentivize more diagnoses per patient. Serden, Lindqvist, and Rosen (2003) look at the introduction of PPS systems county-by-county in Sweden. They found that counties which adopted PPS saw faster growth in diagnoses per patient compared to counties that did not adopt PPS during their sample period.

Between 1983 and 2001, Medicare had a transitional period during which they continued to reimburse retrospectively for capital inputs. This created incentives for hospitals to rely more heavily on capital instead of labor. Acemoglu and Finkelstein (2008) investigated empirically and found evidence that hospitals actually did favor capital during that transitional time period. Their paper took a more short-run approach, using variation across hospitals in the share of Medicare patients as the main source of variation. In contrast, my paper takes a more long-run approach and uses variation across primary diagnosis as the main source of variation. I assume that each hospital interacts with the same technological frontier in the long run. Variation across hospitals comes from the unique relationship each hospital has with that frontier, depending on how fast new technology diffuses to that hospital. My approach assumes that innovation for each primary diagnosis depends on the number of Medicare and privately insured patients with that illness.

The theoretical framework of this paper is most similar to McClellan (1996). McClellan outlines a model where doctors and hospitals have conflicting incentives under prospective payment. McClellan cites case studies in support of his theoretical framework. My model differs from his in that it focuses on R&D rather than cost growth.

2.3 Background on the Relationships between Doctors, Hospitals and Insurers

The relationship between doctors and hospitals is not simple. By law, doctors cannot work for hospitals directly, except as administrators. This is why hospital patients receive two bills: one bill is from the hospital and the other from the doctors. Even in HMO systems where doctors are on salary, the doctors and hospitals operate separately. This is why Kaiser (the

hospital) has a sister organization, Permanente (the doctors' group), to employ the doctors. Kaiser is non-profit, and Permanente is for-profit.

Hospitals either grant or deny privileges to doctors. Hospital privileges allow the doctor to use the hospital's resources, including skilled nursing, labs, technology, operating rooms, and technicians. Doctors want privileges at the hospitals with the best and most readily available resources. Hospitals rely on doctors to bring in patients. Hospitals compete for doctors who bring in the highest quality patients, such as patients on good private insurance plans. If a hospital's resources are already used at capacity by doctors who treat mostly privately insured patients, then that hospital will deny privileges to doctors whose patient mix includes more Medicare and Medicaid patients. The hospital's stock of technology plays a role in attracting doctors with lucrative patients.

Hospitals have several means for influencing doctor behavior (McClellan, 1996). The amount and type of technology a hospital owns is a major way to influence doctor treatment decisions. If a hospital chooses not to purchase a new technology, the doctor has a decision to make. The doctor will either have to send her patient to another doctor who has privileges at a hospital with that technology, or else she will have to treat their patient under the resource constraints of the hospital where she does have privileges.

Of course, hospitals have other means of influencing doctors. Hospitals can deny privileges to doctors who behave in ways they do not like, although they will also lose that doctor's patient base. Guidelines associated with use of the hospital can impact doctor behavior, though only so far as they are enforceable. The hospitals can invest in educational conferences or seminars to teach doctors cost-saving techniques. The hospital has a role in selecting which doctors hold influential positions, such as head of department. They can select doctors who promote specific cost-saving techniques. The hospital can also conduct and publish quality reviews as a way of influencing physician behavior. For example, a hospital might check if doctors are washing their hands before surgery.

The hospital does not have direct control over doctor decision-making. The doctor is entrusted with the responsibility of maximizing the patient's utility. One primary purpose of keeping the doctors and hospitals legally separate is to preserve the freedom of the doctor to make decisions in the patient's best interest.

Doctors are basically paid fee-for-service, both from Medicare and private insurers. They usually make more money for doing surgery than for seeing a patient. Medical ethics may also motivate doctor decisions about treatment intensity. Overall, doctors have little incentive to consider the hospital's cost of treatment, and very often do not even know what various treatments cost. These lead to the conundrum that doctors often recommend a level of treatment such that marginal benefit equals zero, instead of marginal benefit equaling marginal cost.

2.4 Theoretical Model

2.4.1 Overview of Model

There is a single hospital in this economy, and that hospital chooses the macroeconomic level of medical R&D investment. The hospital does not get to determine price because Medicare and the private insurers set the price. The hospital does not determine the quantity directly, because the doctor sets the quantity after observing the level of technology the hospital has chosen. However, hospitals can anticipate that doctors will increase the number of procedures per patient whenever a new technology enters the system.

The percent of patients on Medicare is exogenous. By increasing the number or percent of Medicare patients, we can observe the influence of Medicare's prospective payment system on R&D investment. Essentially, the hospital would solve this problem separately for each illness, plugging in that illness' unique number of elderly and non-elderly patients.

The main result of the model shows that increasing the number of Medicare patients will cause a decrease in R&D investment for therapeutic technologies and an increase in investment for diagnostic technologies. The prospective payment system drives this result. Under such a system, the only way a hospital can increase Medicare revenue is by increasing the number of diagnoses. Any therapeutic treatment will decrease profit at the margin. The model ignores competition between hospitals as a motivating factor for technology adoption. Since investment in technology is the only way for hospitals to influence physician supply decisions, investment will likely follow the incentives set up by the payment scheme. Table 2.1 gives examples of diagnostic and therapeutic procedures. Table 2.1 summarizes the main theoretical predictions of the model. Table 2.2 also adds the prediction in terms of Medicaid patients, assuming that the Medicaid population is representative of charity care patients and that Medicaid reimbursement rates are generally low.

2.4.2 Hospital's maximization problem

The hospital's maximization problem represents how a rational hospital board would think about the decisions regarding how to invest funds for new technologies. The price is set exogenously by Medicare and the private insurers. Doctors determine the quantity. Therefore, the only action left for the hospital to take is to invest in R&D. The hospital chooses the amount of diagnostic and therapeutic technology (τ_D and τ_T) to develop. The hospital makes this decision with the knowledge that an increase in technology will lead the doctors to increase the number of procedures they perform. The doctor will choose to perform a higher number of procedures per patient when there is more technology available, and this is true whether the doctor is motivated by maximizing the patient's utility or her own income. For simplicity, each new unit of technology will lead to one additional procedure per patient: $\frac{\partial proc}{\partial \tau} = 1$.

Each new diagnostic procedure performed will lead to a greater number of diagnoses, but

Diagnostic Procedures	Therapeutic Procedures
IMAGING	SURGERY
CT scan of brain	Cardiac catheterization
MRI of abdomen	Insertion of coronary stent
Ultrasound of fetus	Angioplasty using laser
Angiography, femoral	Abdominal hysterectomy
	Removal of appendix
BIOPSY	
Pap smear	OTHER
Biopsy of liver	Blood transfusion
Biopsy of lung tissue	Radiation therapy
	Mechanical ventilation
OTHER	Kidney dialysis
Hearing test	Electrocoagulation of uterus
Measurement of arterial gases	Injection of drugs into spine

Table 2.1: Examples of diagnostic and therapeutic procedures

An increase in	leads to this change in technological investment
Medicare population	Decrease in therapeutic technology
Medicare population	Decrease or increase in diagnostic technology
Privately insured population	Increase in therapeutic technology
Privately insured population	Increase in diagnostic technology
MediCal/Medicaid population	Decrease in therapeutic technology
MediCal/Medicaid population	Decrease in diagnostic technology

Table 2.2: Summary of predictions from the theoretical model

there are diminishing marginal returns to additional diagnostic procedures: $\frac{\partial diag}{\partial proc_D} > 0$ and $\frac{\partial^2 diag}{\partial proc_D^2} < 0$.

The hospital incurs investment costs, $I(\tau_D, \tau_T)$, which depend on how much technology it chooses to develop. This hospital is investing directly in such technology, and is therefore responsible for hiring the scientists and buying the research equipment. Because of scarce resources, investment costs increase at an increasing marginal rate as the hospital produces a greater number of technologies, so $\frac{\partial I}{\partial \tau_j} > 0$ and $\frac{\partial^2 I}{\partial \tau_j^2} > 0$ where $j \in (D, T)$.

Also, diagnostic and therapeutic technologies are neither perfect substitutes nor perfect compliments in terms of their impact on total investment costs. For a fixed value of I , the properties of $I(\tau_D, \tau_T)$ are such that the marginal rate of transformation between τ_D and τ_T is as follows: $\frac{\partial \tau_T}{\partial \tau_D} < 0$, $\frac{\partial^2 \tau_T}{\partial \tau_D^2} < 0$, $\frac{\partial \tau_D}{\partial \tau_T} < 0$, and $\frac{\partial^2 \tau_D}{\partial \tau_T^2} < 0$.

For simplicity, the hospital's production cost is simply a fixed cost per procedure, C .

Medicare revenue will be equal to the number of Medicare patients, pop_k^{MC} , times the reimbursement rate per diagnosis, R^{MC} , times the number of diagnosis per patient, $diag$. Private insurer revenue is equal to the number of privately insured patients, pop_k^{priv} , times the reimbursement rate per procedure, R^{priv} , times the total number of procedures per patient, $proc_D + proc_T$. These two different setups reflect the differences between Medicare's prospective payment and the fee-for-service payment setup that private insurers generally use for hospitals.

For every primary diagnosis, k , the hospital will solve a profit maximization problem to determine the level of investment in technology that treats patients with that particular illness:

Profit = Medicare revenue + Private insurer revenue – Production costs – Investment costs

$$\begin{aligned}
\Pi &= op_k^{MC} R^{MC} \cdot diag [proc_D(\tau_D)] \\
&\quad + pop_k^{priv} R^{priv} [proc_D(\tau_D) + proc_T(\tau_T)] \\
&\quad + (pop_k^{MC} + pop_k^{priv}) \cdot C \cdot [proc_D(\tau_D) + proc_T(\tau_T)] \\
&\quad - (pop_k^{MC} + pop_k^{priv}) \cdot C \cdot [proc_D(\tau_D) + proc_T(\tau_T)] \\
&\quad - I(\tau_D, \tau_T)
\end{aligned} \tag{2.1}$$

where

pop_k^{MC} = number of people with primary diagnosis k who are on Medicare

pop_k^{priv} = number of people with primary diagnosis k who are privately insured

$pop_k^{total} = pop_k^{MC} + pop_k^{priv}$

R^{MC} = Medicare's fixed dollar payment per diagnosis

R^{priv} = private insurers' fixed dollar payment per procedure

C = fixed cost to hospital of producing one procedure

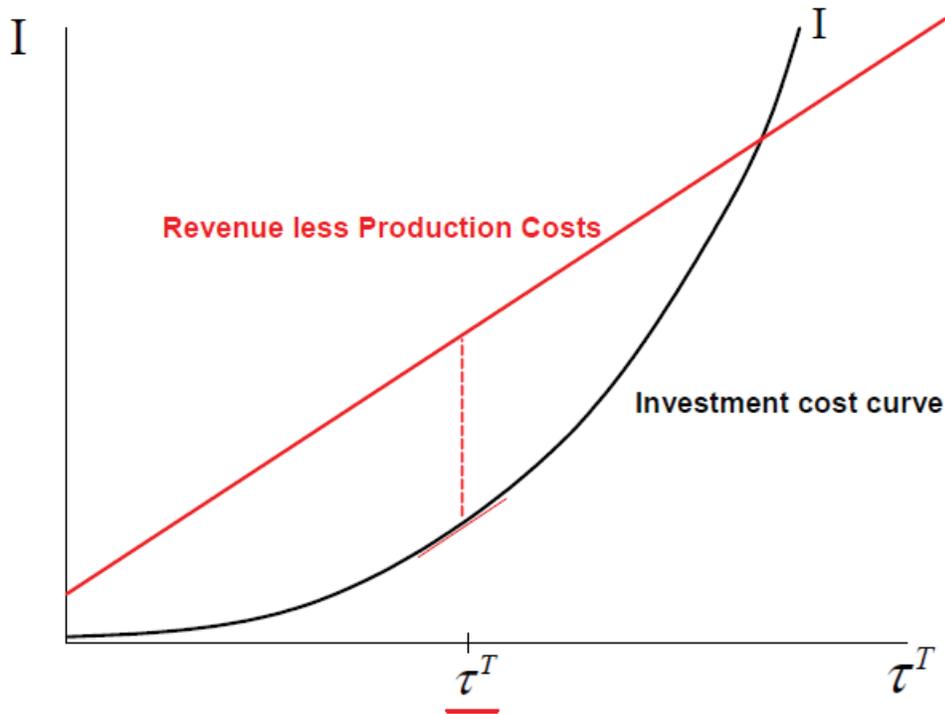


Figure 2.1: Graphical representation of the hospital's maximization problem (Equation 2.1)

$diag$ = diagnosis per patient

$proc_j$ = procedures per patient, $j \in (D, T)$

τ_j = units of technology, $j \in (D, T)$

I = investment costs

D = diagnostic

T = therapeutic

Figure 2.1 shows a graphical representation of Equation 2.1. Because of the simplifying assumptions, investment is the only source of curvature. This curvature arises from the fact that there are scarce R&D resources.

The hospital will solve this problem separately for each diagnostic group, k . The hospital will choose one level of R&D investment for technologies aimed at breast cancer, another level of R&D investment for technologies aimed at brain strokes, etc. This allows for the primary diagnosis to be the main source of variation for empirical investigation. Figure 2.2 shows how the model supports a different level of investment depending on how many Medicare patients have that illness.

Only one hospital exists in this model, and that hospital invests directly in R&D. However, the model can be extended to include more hospitals and a separate R&D industry with

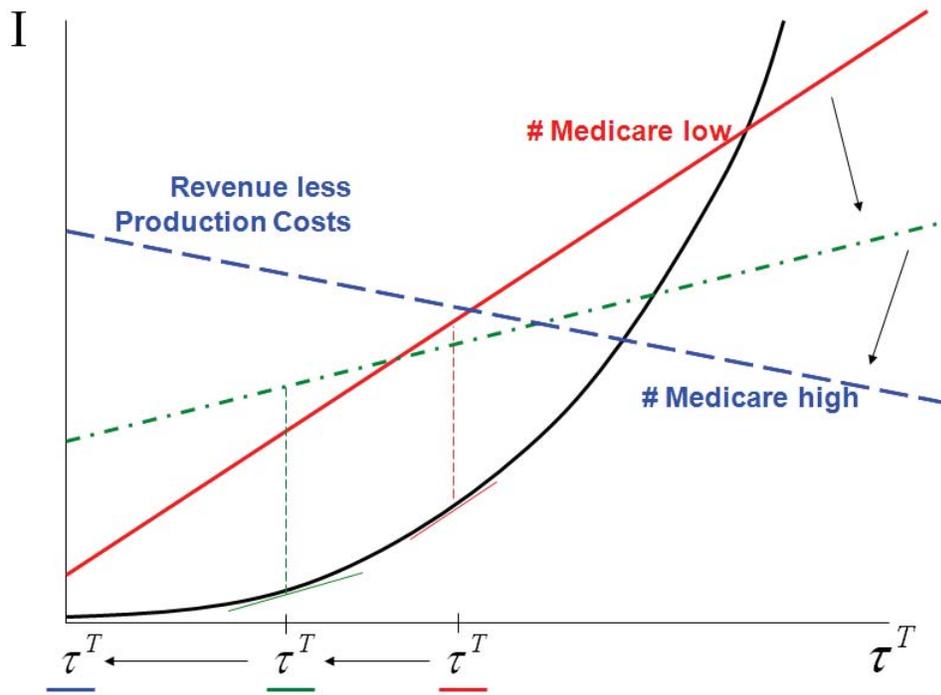


Figure 2.2: Medicare patients as source of variation
(Equation 2.1 with different primary diagnosis)

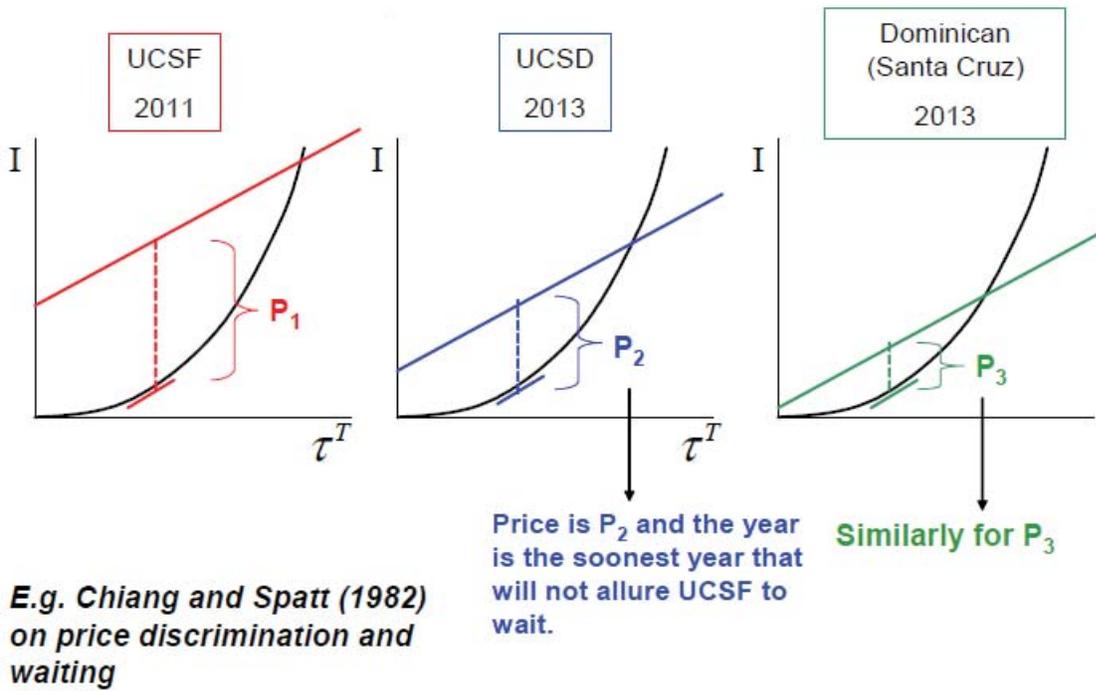


Figure 2.3: Price discrimination through waiting

very little disruption to the model's current framework. For one, the percent of elderly with a particular illness should be similar across most hospitals. Larger and more prestigious hospitals may still have a higher demand for those technologies. However, from a profit maximization standpoint, it would be sufficient for the R&D firms to choose the level of R&D investment that the largest and most prestigious hospital would choose if it owned all R&D. Then the R&D firms could sell the same technologies to the other hospitals at lower prices after different wait periods. This follows the price discrimination by waiting model (Chiang and Spatt, 1982), which involves a diffusion process for new technologies. Figure 2.3 demonstrates these concepts graphically.

2.4.3 Solving the Model

To solve the model, take the first order conditions of Equation 1:

$$\frac{\partial \Pi}{\partial \tau_T} : pop_k^{priv} R^{priv} - (pop_k^{MC} + pop_k^{priv}) \cdot C - \frac{\partial I}{\partial \tau_T} = 0 \quad (2.2)$$

$$\begin{aligned} \frac{\partial \Pi}{\partial \tau_D} : & \text{pop}_k^{\text{priv}} \cdot R^{\text{priv}} \frac{\partial \text{diag}}{\partial \text{proc}_D} + \text{pop}_k^{\text{priv}} R^{\text{priv}} \\ & - \left(\text{pop}_k^{\text{MC}} + \text{pop}_k^{\text{priv}} \right) \cdot C - \frac{\partial I}{\partial \tau_T} = 0 \end{aligned} \quad (2.3)$$

Equation 2 shows us how investment in therapeutic technologies will change in response to a change in the number of Medicare patients or privately insured patients. Given that an increase in τ_T will lead to an increase in $\frac{\partial I}{\partial \tau_T}$, we can see that therapeutic investment will decrease in response to a higher number of Medicare patients, and that it will increase in response to a higher number of privately insured patients.

Equation 3 reveals that diagnostic investment will increase in response to privately insured patients. This stems from the fact that private insurer reimbursement rates must be higher than the cost of treatment, C , in order for the hospital to accept that insurance. However an increase in the number of Medicare patients could actually increase diagnostic investment, although the exact prediction is ambiguous. It depends on Medicare's reimbursement rate, R^{MC} , and the return in new diagnosis of each diagnostic procedure, $\frac{\partial \text{diag}}{\partial \text{proc}_D}$. However, it is very plausible that an increase in Medicare patients would increase diagnostic investment, so long as reimbursement rates are sufficient. Diagnostic procedures are not the main focus of this paper. Rather, diagnostic procedures act as a control group for comparison. Even if there is a negative relationship between the number of Medicare patients and investment in diagnostic technology, that negative relationship should be less dramatic than for therapeutic investments.

2.5 Empirical Strategy

2.5.1 Data

The analysis employs two different data sets to test the theoretical predictions. The OSHPD data set contains a record of every patient in the state of California for the years 2001 to 2008. It is a high-quality data set that includes hospital identifiers, up to 25 diagnoses and 21 procedures per patient and detailed patient characteristics. On the other hand, the National Hospital Discharge Survey is of lower quality, but provides data on most years between 1980 and 2006. Comparing the years of overlap tests for the validity of the lower quality historical data set.

In some specific areas, the historical data set falls short. The Discharge Survey does not indicate hospital, so it is impossible to implement hospital fixed effects. As a proxy, I create hospital groups based on region, hospital size, and ownership type. The historical survey also only allows hospitals to list up to 7 diagnoses and 4 procedures, while the OSHPD data allows for 24 diagnoses and 20 procedures. The historical survey data is subject to sampling

bias, while OSHPD data includes all hospitals in California. Inconsistencies over time may arise from the historical survey, because the sampling methodology was updated in 1988, and selection into the sample was updated in 1991, 1994, 1997, 2000, and 2003.

The biggest problem with the historical data was that the survey does not specify which diagnosis is the primary diagnosis, which is a key piece of information necessary for my analysis. I address this problem in two ways. First I did the analysis assuming that the first diagnosis listed was the primary diagnosis. Second, I threw out all patients with more than one diagnosis, and ran the analysis on this subset. Although each of these approaches has disadvantages, both analyses told the same story.

2.5.2 Empirical Model

One problem with investigating R&D incentives is that there is only one technological frontier interacting with all hospitals. Each hospital has a unique relationship to that technological frontier, depending on how quickly new technology diffuses to that particular hospital. But each hospital does not have its own technological frontier. For this reason, comparison across hospitals will not work as an empirical approach.

In order to elicit variation, this empirical model treats the technological frontier associated with each primary diagnosis as unique and independent. For example, innovation geared at breast cancer makes up a separate technological frontier than innovation geared at cardiac arrest or brain strokes.

I use the number of procedures performed on a patient as an indication of the accumulated technology geared at that patient's primary diagnosis. Of course, new technology could potentially decrease the number of procedures per patient. However, given the extremely rapid rise in procedures per patient over time, it seems that, on average, new technologies are adding to the number of procedures rather than displacing procedures. For example, between 2001 and 2008, there was an eight percent increase in procedures per patient.

The main regression tries to answer this question: Does a patient with an illness common among the elderly receive fewer procedures than a patient with an illness common among the young? This means that in addition to individual-level characteristics, I needed to include characteristics of the patient's diagnostic group. For example, I calculate the number of people with the same diagnosis who are elderly and the number who are non-elderly. These become the key variables of interest. By including the number rather than the percent, I capture the fact that the popularity of an illness will have a big impact on investment dollars, but that popularity among the elderly has the opposite effect of popularity among the non-elderly.

Tables 2.5 and 2.6 show the basic model: Hospital fixed effects, excluding "do not resuscitate" patients, in logs, restricting data to patients ages 60 to 70:

$$\#procJ_{i,k} = \beta_0 + \beta_1\#old_k + \beta_2\#young_k + \beta_3\#MediCal_k$$

$$+\beta_4 E_k + \beta_5 F_i + \varepsilon_{i,k} \quad (2.4)$$

where

i = individual patient

k = primary diagnosis

$\#proc_{J,i,k}$ = number of diagnostic or therapeutic procedures performed on patient i with primary diagnosis k

$J \in (D, T)$ is diagnostic or therapeutic

$\#old_k$ = total number of elderly patients with primary diagnosis k

$\#young_k$ = total number of non-elderly patients with primary diagnosis k

$\#MediCal_k$ = total number of MediCal patients with primary diagnosis k

E = a vector of controls for primary diagnosis, k . This includes the average age for elderly patients with diagnosis k and the average age for young patients with primary diagnosis k . When regressing on diagnostic procedures, the number of diagnoses for the average patient in diagnostic group k is also included in E . This represents the fact that it will be more worthwhile for doctors to perform diagnostic procedures on patients with illnesses that are associated with a higher number of co-morbidities.

F = a vector of controls for individual patient characteristics. This includes age, sex, and a dummy for whether the patient is elderly. When regressing on therapeutic procedures, the number of diagnoses for each individual patient is also included in F

Age differences in treatment intensity posed a particular problem. Figure 2.4 shows that procedures per patient roughly increase until sometime around retirement, and then decrease thereafter. There is also a strong impact during the child-bearing years. Interestingly enough, the pattern observed during child-bearing years fits the story very well. Namely, there are more therapeutic procedures performed on pregnancy-related conditions. Since pregnancy will almost never be covered by Medicare, the disincentive for development of therapeutic procedures simply is not there.

The hill shape observed in Figure 2.4 made it difficult to control for age directly. My solution was to restrict the analysis to patients from ages 60 to 70, where there was no statistically significant relationship between age and number of procedures, as Figure 2.5 demonstrates graphically. I calculated the k-level variables (diagnostic group) with the full age spectrum, and but excluded people under 60 and over 70 when running the regressions (Tables 2.5 and 2.6).

Of course, the age distribution of a particular illness may impact technological development independently from the payment systems. To control for this, I included separately the average age for elderly with diagnosis k and the average age for non-elderly. The necessity of having these two separate measures arises from the reversal in the direction of the age-procedure relationship that happens sometime around age 65. Based on the theoretical expectations outlined in Table 2.2, we would expect a higher average age among the non-elderly to be associated with more procedures, and a higher age among the elderly to lead

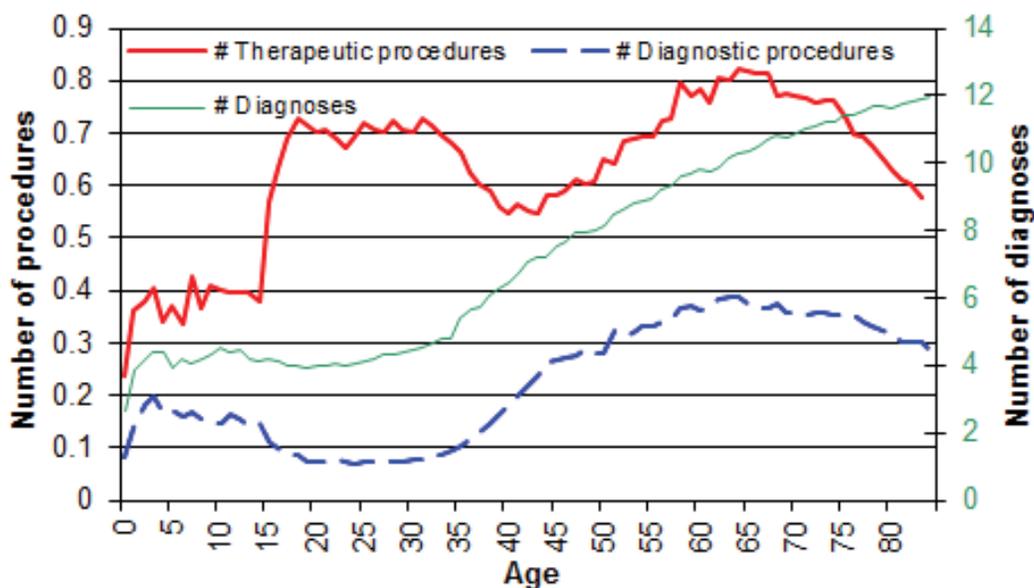


Figure 2.4: Number of Procedures and Diagnoses Per Patient by Age
(The reason to limit analysis to patients age 60 to 70)

to fewer procedures. This is exactly what the data shows. The elderly dummy builds in a regression discontinuity that will help control for differential treatment of elderly patients.

The number of patient diagnoses is the only variable that is different in the diagnostic regression compared to the therapeutic regression. We would expect that the patient will undergo diagnostic procedures before therapeutic procedures. Therefore, the number of therapeutic procedures will depend on the patient's actual number of diagnoses. On the other hand, the number of diagnostic procedures will depend on the expected number of diagnoses, as predicted by their primary diagnosis.

Fixed effects control for hospital-level differences in treatment intensity. Each hospital has a unique relationship with the technological frontier since some hospitals are early adopters and other late adopters. Tables 2.3 and 2.4 shows examples of why fixed effects are helpful and why I wanted to run the model without fixed effects as a robustness check. Some hospitals perform more procedures on all types of patients (like UCSF), likely because they adopt technology more quickly than other hospitals.

I excluded observations where the patients received a “do not resuscitate” directive within 24 hours of entering the hospital. This accounted for 4.7% of patients. For the diagnoses I have coded, I excluded patients whose primary diagnosis was either psychological or substance-abuse related. This accounted for about 8% of patients. Since psychological and substance-abuse patients are primarily treated through psychotherapy, drugs and supervision, the number of procedures seemed less relevant as a proxy for technological innovation

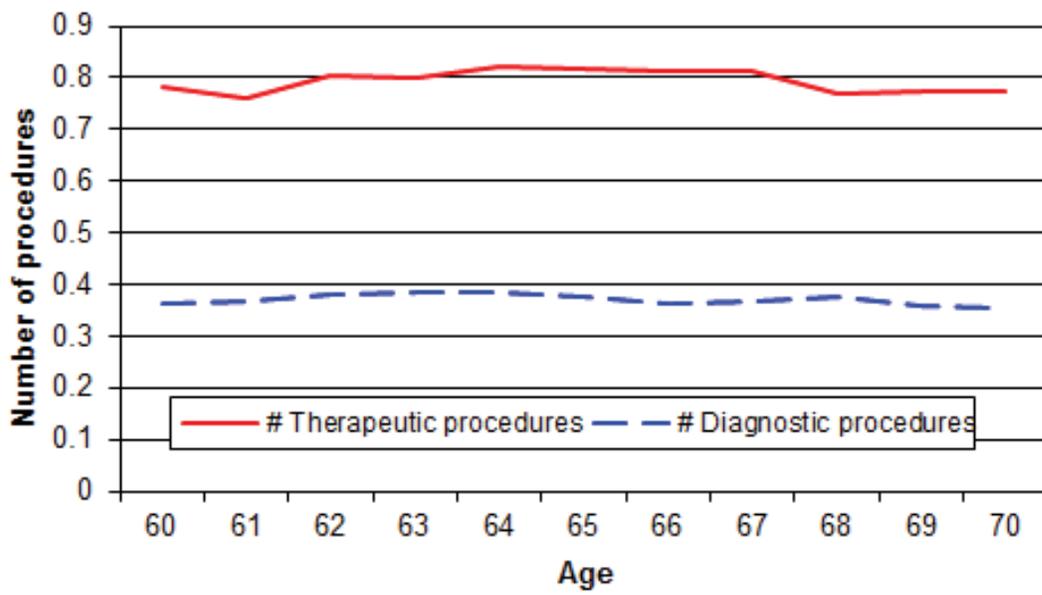


Figure 2.5: Procedures Per Patient by Age – Restricted to ages 60 to 70
(There is no significant relationship between age and number of procedures for patients age 60 to 70)

# therapeutic procedures for patient with...	Univ of CA - San Francisco	Univ of CA - San Diego	Dominican (Santa Cruz)	Santa Rosa Memorial
Cancer of the ovaries	4.2	4.3	1.6	1.6
Osteoarthritis (joint disorder)	1.14	1.09	0.42	0.39
Ruptured brain artery	2.24	2.00	1.25	0.55

Table 2.3: Reason for using hospital fixed effects
Examples of hospitals. Across different classes of hospitals, there are patterns in the number of procedures.

# therapeutic procedures for patient with...	Dominican	Mercy Medical	Ventura County Medical	St. John's Regional
Cancer of the ovaries	1.6	1.4	1.8	3
Osteoarthritis (joint disorder)	0.42	0.36	0.38	0.55
Ruptured brain artery	1.25	1.6	2.25	1.0

Table 2.4: Reason for checking analysis without using hospital fixed effects Examples of hospitals.

geared at this type of patient.

The addition of the variable for number of MediCal patients allows for testing of a third payment system. MediCal is California’s version of Medicaid. Nationwide each state determines its own payment scheme for Medicaid. However, Medicaid payments are known to be reimbursed at a level significantly below that of private insurers in most states. Also MediCal’s patient base will resemble the demographic composition of patients who receive charity care. The MediCal variable actually includes three categories of patients: (1) twenty-five percent of patients are on MediCal, (2) two percent of patients are in the “county indigent” category and (3) 0.3 percent of patients are in the “other indigent” category. Because of this, we would expect both fewer diagnostic and fewer therapeutic procedures performed on patients with illnesses common among MediCal patients.

Inclusion of a variable for the number of MediCal patients helps test the strength of the model. Based on the hospital payment setup surrounding the demographic of people covered by MediCal, we expect both fewer diagnostic procedures and therapeutic procedures. This hypothesis differs from our expectation for both the non-elderly population and the elderly population bases. If the results support our predictions for all three demographic groups, it is more likely that this empirical model accurately captures differences in technology innovation driven by the hospital payment systems associated with each group.

2.6 Results

Figures 2.6, 2.7, 2.8 and 2.9 show simple summary statistics of the number of procedures per patient for illnesses with different shares of elderly and by different age brackets. The pattern that emerges is the one expected based on the incentives of Medicare’s prospective payment system. Namely, Figure 2.6 indicates that illnesses with a higher share of elderly have fewer therapeutic procedures per patient. Figure 2.7 shows a similar negative relationship regarding the percent of patients in the MediCal/indigent care group. This relationship is absent when looking at diagnostic procedures (Figure 2.8), and also disappears when looking at 1983 data (Figure 2.9).

The regression results in Table 2.5 also support the theoretical predictions summarized in Table 2.2. Specifically, illnesses common among the elderly are associated with fewer therapeutic treatments. Indeed, there is a negative correlation between the number of elderly with a patient’s primary diagnosis and the number of therapeutic procedures performed on that patient. This makes sense in light of a system that pays hospitals by the diagnoses, with nothing extra for specific procedures.

On the other hand, since Medicare’s prospective payment system pays by diagnoses, there is an incentive to develop more diagnostic procedures. So we expect a positive relationship between the number of elderly and diagnostic procedures. As expected, the regression in Table 2.5 yields a positive coefficient when using diagnostic procedures as the dependent variable. Since private insurers generally reimburse based on procedures – both diagnostic

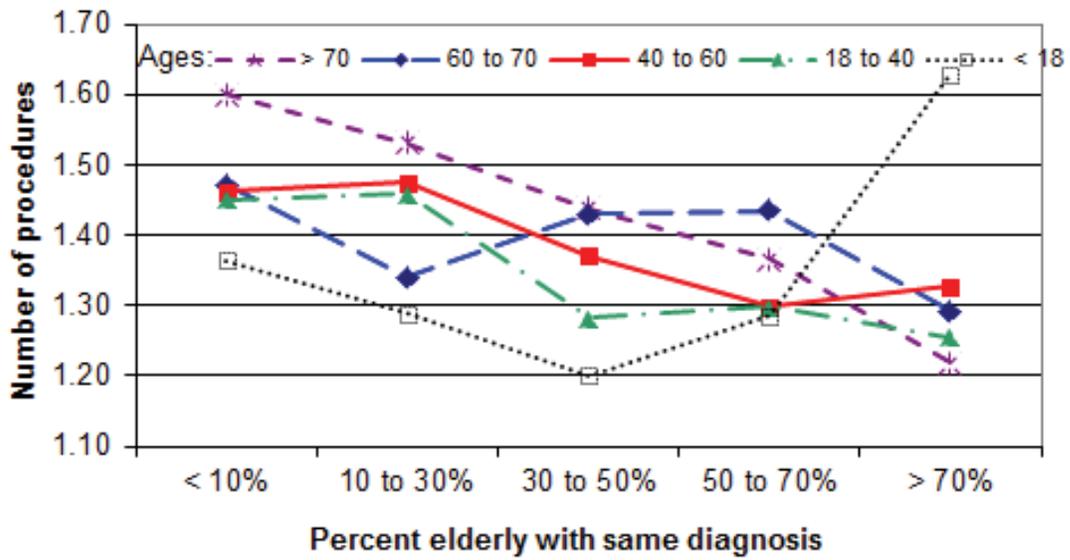


Figure 2.6: Average number of therapeutic procedures per patient
 For diagnoses common and uncommon among the elderly, by age group. 2008 OSHPD data.
 (Summary statistics that preview the regression results)

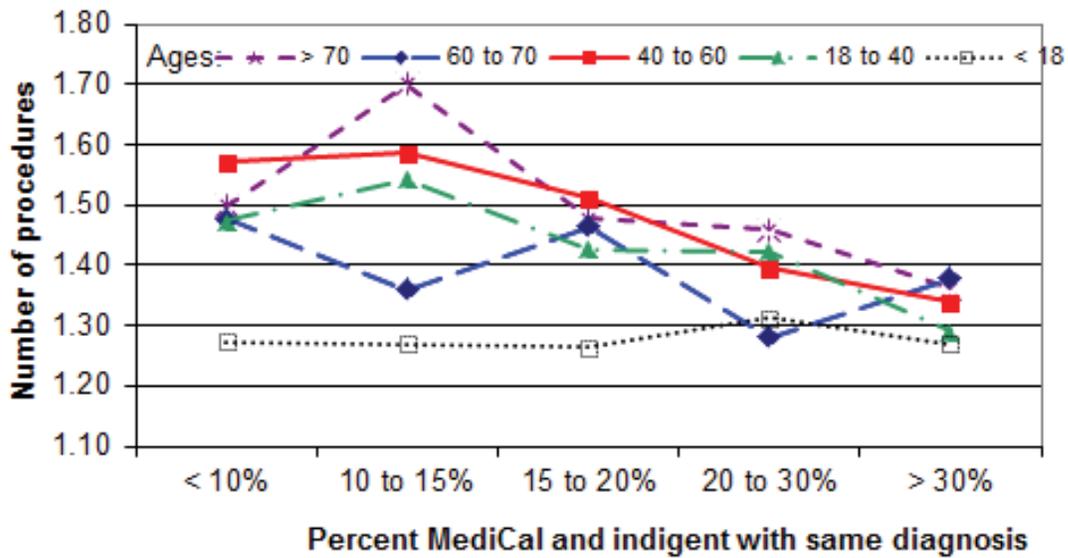


Figure 2.7: Average number of therapeutic procedures per patient For diagnoses common and uncommon among MediCal and indigent patients, by age group. 2008 OSHPD data. (Summary statistics that preview the regression results)

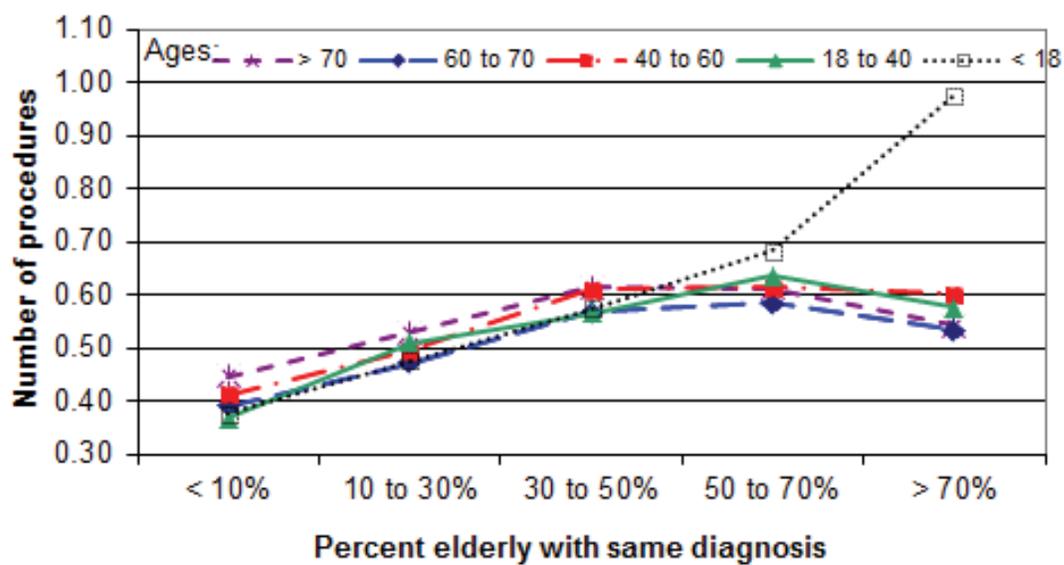


Figure 2.8: Average number of diagnostic procedures per patient
 For diagnoses common and uncommon among the elderly, by age group. 2008 OSHPD data.
 (Summary statistics that preview the regression results)

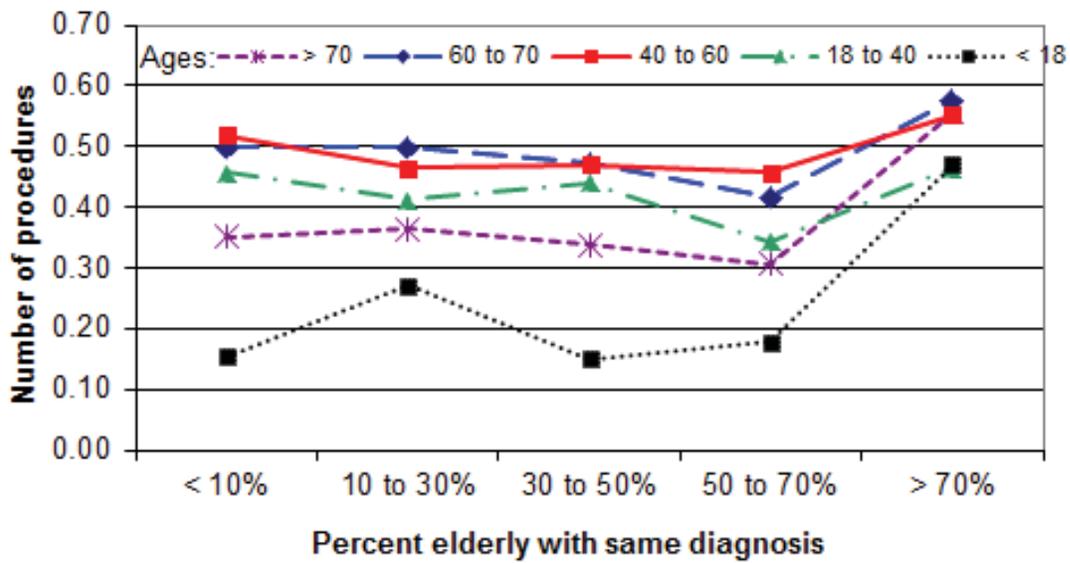


Figure 2.9: Average number of therapeutic procedures per patient
 For diagnoses common and uncommon among the elderly, by age group. 1983 Survey data.
 (The relationship from Figure 2.6 is not present in 1983 data)

		Diagnostic (All patients)	Diagnostic (1 diag patients)	Therapeutic (All patients)	Therapeutic (1 diag patients)
# Elderly patients in diagnostic group	k	0.013***	0.076*	-0.036***	-0.081***
# Young patients in diagnostic group	k	0.027***	-0.115	0.122***	0.055**
# Medicaid patients in diagnostic group	k	-0.026***	0.054	-0.104***	-0.011
Mean age of elderly	k	-0.523***	-0.998	-0.178**	-1.460***
Mean age of non-elderly	k	0.201***	-0.239	0.326***	0.526***
Mean # diagnosis per patient	i	0.113***	0.335***		
# Diagnosis for this patient	i			0.304***	(dropped)
Female	i	-0.029***	0.015	-0.010***	0.022
Elderly	i	-0.009	-0.106	0.005	0.022
Age	i	0.108	1.236	-0.237***	-0.080
Constant		1.052*	-0.278	0.391	5.084*
N (i-level)		59,769	382	121,425	1,795
N (k-level)		2,437	183	2,989	426

Table 2.5: (2008) Effect of elderly on procedures
Relationship between the number of elderly in a patient's diagnostic group and the number of procedures performed
on that patient. Regressing on number of procedures (either diagnostic or therapeutic) per patient. 2008 OSHPD
data.

		Diagnostic (All patients)	Diagnostic (1 diag patients)	Therapeutic (All patients)	Therapeutic (1 diag patients)
# Elderly patients in diagnostic group	k	0.017*	0.024*	0.035***	0.103***
# Young patients in diagnostic group	k	-0.033**	-0.069***	-0.008	-0.070***
# Medicaid patients in diagnostic group	k	0.015	0.036*	-0.020***	-0.005
Mean age of elderly	k	-0.048	0.051	-0.197	-0.037
Mean age of non-elderly	k	-0.047*	-0.010	0.004	-0.068***
Mean # diagnosis per patient	i	0.149***	0.046		
# Diagnosis for this patient	i			0.070***	(dropped)
Female	i	0.014	-0.014	0.014*	0.020
Elderly	i	0.016	0.029	0.023	0.003
Age	i	-0.148	-0.462	-0.134	0.160
Constant		1.158	2.045	1.488	-0.193
N (i-level)		6,228	1,287	8,793	2,853
N (k-level)		946	439	1,193	665

Table 2.6: (1983) Effect of elderly on procedures
Relationship between the number of elderly in a patient's diagnostic group and the number of procedures performed on that patient. Regressing on number of procedures (either diagnostic or therapeutic) per patient. 1983 National Discharge Survey.

and therapeutic – we expect a positive correlation between the number of non-elderly with an illness and procedures targeting that illness. That is exactly what we observe.

Additionally, illnesses common among MediCal patients are associated with both fewer diagnostic and therapeutic treatments. This is true, even when restricting the sample to privately insured patients between the ages of 60 and 64. In a system where low-income patients must be treated with charity care, you would expect development of fewer technologies aimed at low-income patients.

These results bolster confidence that even privately insured patients are impacted by payment systems of Medicare and Medicaid. When a patient shares a diagnosis with elderly or low income patients, that patient’s treatment options will be impacted by the hospital’s average profitability for their illness. And if that profitability relies heavily on Medicaid or Medical, fewer therapeutic technologies will be developed.

Of course, Table 2.5 only shows a snapshot in time. Ideally we would like to compare it to the counterfactual of a world without prospective payment. Luckily, the historical data set brings us back to a time before Medicare implemented prospective payment. Table 2.6 shows the regression results using the 1983 data. Medicare implemented prospective payment in October of 1983. Figures 2.10, 2.11 and 2.12 plot the key coefficient from Tables 2.5 and 2.6 over time.

Given the theory and intuition outlined in this paper, what did we expect from the historical data? The hypothesis is that Medicare’s prospective payment is causing the negative relationship between number of therapeutic procedures per patient and the number of elderly with the patient’s primary diagnosis. If the relationship is causal, this negative relationship should disappear when looking at data prior to 1983, the year Medicare implemented prospective payment. Specifically, the coefficient on $\#old_k$ (number of elderly by primary diagnosis) should be positive or insignificant prior to 1983. This coefficient should decline after 1983 at a pace that depends on how fast new technology displaces old technology.

Figures 2.10, 2.11, and 2.12 generally support these predictions. In Figure 2.10, the coefficient on the number of elderly peaks in 1984, the year after Medicare implemented prospective payment. In years following 1984, the coefficient declines, with an increase in the decline in the years leading up to the 2001 final transition to prospective payment for capital inputs. Figure 2.11 shows an increase in the coefficient in the mid-1990’s that may be resulting from the fact that inclusion of all patients requires me to invoke the unlikely assumption that the first diagnosis listed is the primary diagnosis. There were no instructions on the survey requiring hospitals to list the primary diagnosis first, so the accuracy of these results rests on the hopes that hospitals would list them in descending order of importance naturally. Restricting the sample to patients with one diagnosis, as in Figure 2.10, makes more sense for data from the early 1980’s when nearly a third of patients had only one diagnosis. This restriction becomes more problematic over time as fewer and fewer patients fall into that category. By 2001, only 1.5 percent of patients have a single diagnosis. Luckily, the historical data is most useful for its predictions from the early 1980’s, and we can take the Survey Data results from the 1990s and 2000’s with a grain of salt. Thankfully the two

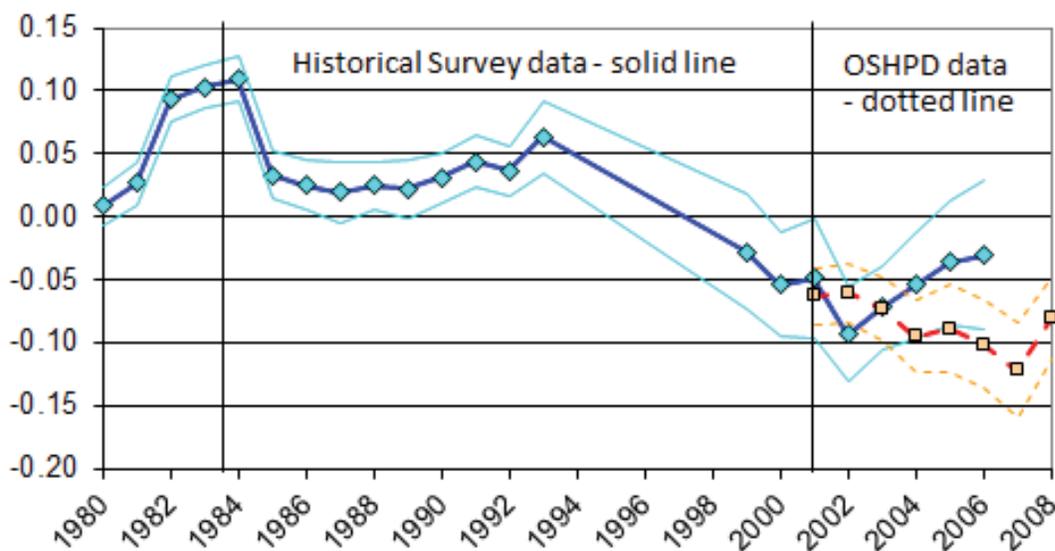


Figure 2.10: Coefficient on number of elderly

Coefficient on number of elderly when regressing the log number of therapeutic procedures on the log of the number of elderly patients sharing this patient's primary diagnosis, plotted with 95% confidence intervals for 1980 to 2006 (Patients with one diagnosis: fourth column from Tables 2.5 and 2.6) Special dates marked with a vertical line: October 1983 marks the implementation of the prospective payment system. In order to give hospitals time to make capital adjustment, the system allowed for a transition period during which hospitals still received retrospective payment for capital inputs. This transition ended in 2001, but was widely anticipated since the 1980's.

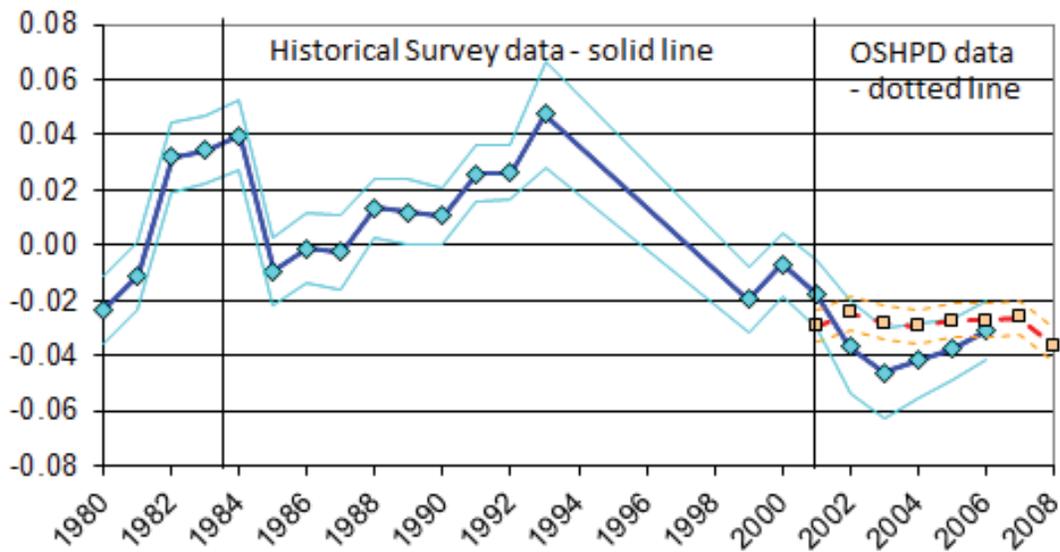


Figure 2.11: Coefficient on number of elderly

Coefficient on number of elderly when regressing the log number of therapeutic procedures on the log of the number of elderly patients sharing this patient's primary diagnosis, plotted with 95% confidence intervals for 1980 to 2006 (All patients: third column from Tables 2.5 and 2.6).

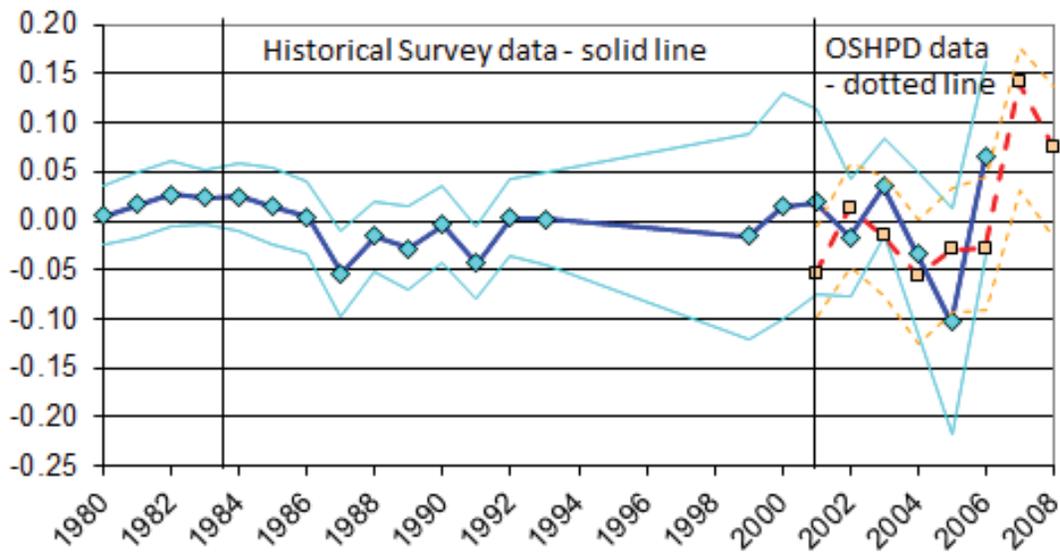


Figure 2.12: Coefficient on number of elderly
 Coefficient on number of elderly when regressing the log number of diagnostic procedures on the log of the number of elderly patients sharing this patient's primary diagnosis, plotted with 95% confidence intervals for 1980 to 2006 (Patients with one diagnosis: second column from Tables 2.5 and 2.6).

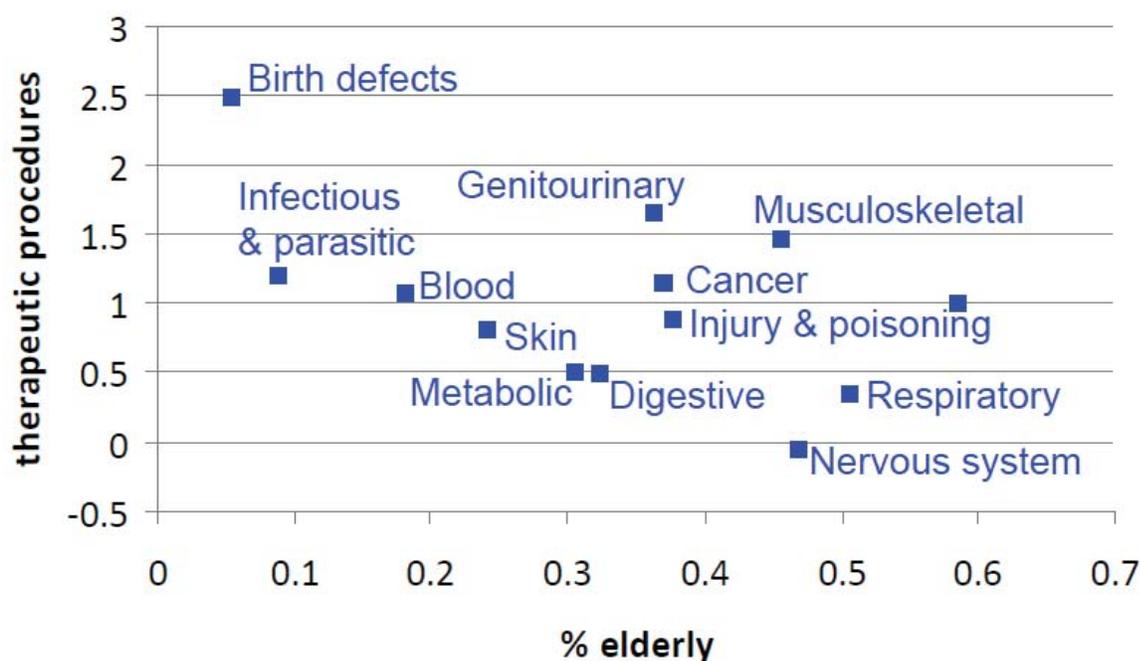


Figure 2.13: Scatter plot of therapeutic procedures v. percent elderly for different categories of illnesses
(Patients ages 60-to-70, controlling for number of diagnosis)

data sets seem to be telling a similar story for the years of overlap.

These results support the notion that Medicare's prospective payment system suppresses development of therapeutic technology. More broadly, this indicates that hospital financial incentives are impacting which technologies innovators are choosing to develop.

2.7 Robustness checks

The diagnostic codes can be grouped into seventeen specific categories, as defined by the ICD-9-CM codes listed for each diagnosis. These categories include cancer, blood-related illnesses, respiratory illnesses, and other broad classifications listed in Figure 2.13 and Table 2.7. If hospital payment schemes are indeed playing a role in channeling R&D dollars, then we should observe the negative relationship between elderly and therapeutic procedures both across categories and within categories.

Figure 2.13 and Table 2.7 show that the expected relationship generally holds both across categories and within categories. Figure 2.13 shows a scatter plot of the number of therapeutic procedures against the percent elderly in each category, controlling for number of diagnoses and restricting to 60-to-70-year-old patients to control for age. The scatter plot

Category	supports theory?	Elderly coefficient	Young coefficient	N (i-level)	N (k-level)
Musculoskeletal	Yes	-0.27***	0.24	14,446	301
Circulatory	Yes	-0.17***	0.15***	25,168	325
Respiratory	Yes	-0.15***	0.14***	8,893	333
Genitourinary	Yes	-0.08***	0.10	6,233	317
Injury and poisoning	Yes	-0.08***	0.10	12,958	331
Infectious and parasitic	Yes	-0.06***	0.09***	7,598	286
Cancer	Yes	-0.05***	0.04***	12,125	311
Nervous system	Yes	-0.04***	0.05**	6,408	321
Digestive	Yes	-0.02*	-0.01	14,715	326
Blood	Neutral	0.00	-0.03**	3,773	324
Skin	No	0.06*	-0.09***	2,098	310
Metabolic & nutritional	No	0.07***	-0.07***	4,222	314
Birth defects	No	0.25***	-0.12**	182	89
Pregnancy-related	N/A	N/A	N/A	N/A	N/A
Perinatal	N/A	N/A	N/A	N/A	N/A

Table 2.7: Regressing within each category.

Relationship between the number of elderly in a patient's diagnostic group and the number of procedures performed on that patient

Specification changes:	Coefficient on elderly in diagnostic group	Coefficient on non-elderly in diagnostic group
Baseline	-0.04***	0.02***
Treatment v. non-treatment (as a test for extensive margin)	0.09***	-0.11***
Percent elderly instead of numbe elderly	-0.02***	n/a
ADJUSTMENT OF CONTROLS		
Excluding all i-level controls	-0.03***	0.02***
Excluding all k-level controls	-0.03***	0.02***
Excluding all variables except eld_k	-0.01***	n/a
Ages 40 +	-0.03***	0.00**
All ages	-0.01***	-0.01***
No hospital fixed effects	-0.03***	0.02***

Table 2.8: 2008 Robustness checks: Different specifications

reveals a clear negative relationship, as the model predicts. Table 2.7 shows the results when running the main regression separately on all diagnoses within a particular category. In nine of the categories, the coefficient on number of elderly is negative, as the model predicts. This coefficient is positive on three of the categories, and insignificant on one. Two of the categories, pregnancy and parental, have no variation in the number elderly across diagnoses within the category, so I was unable to run the regression on those. Both of these analyses generally support the theoretical predictions that there is a negative relationship between elderly and therapeutic procedures across diagnostic groups.

The results seem to be along the intensive margin rather than the extensive margin. Table 2.8 shows that when replacing number of procedures with a dummy for whether or not a patient had any procedures at all, the negative relationship disappears. This is likely due to the fact that the kinds of illnesses that are not generally treated may have characteristics that make them poor candidates for future treatment innovation. These illnesses may be too mild to risk the side effects of additional procedures, or they may be illnesses where the main function the hospital plays is a monitoring function. When I exclude illnesses that are in the bottom 25th percentile and 50th percentiles in terms of how often they are treated (Table 2.9), the regression results get stronger, which bolsters the result that the innovation changes I am measuring tend to be along the intensive margin.

Tables 2.8, 2.9, 2.10, and 2.11 show that the results are not sensitive to the exclusion of outliers, the exclusion of controls and the exclusion of fixed effects. The only time the main coefficient changes signs is when testing the extensive margin, as described above. In all other cases for 2008, there remains a negative and statistically significant coefficient on

Excluding outliers:		Coefficient on elderly in diagnostic group	Coefficient on non-elderly in diagnostic group
No restrictions (Baseline)		-0.04***	0.02***
Avg # therapeutic procedures in group (10th to 90th percentile: 0.23 to 2.32)	k	-0.03***	0.04***
Avg # therapeutic procedures in group (Above 50th percentile: > 1.08)	k	-0.10***	0.09***
Avg # therapeutic procedures in group (Above 25th percentile: > 0.47)	k	-0.06***	0.05***
# Therapeutic procedures for this patient (Below 90th percentile: < 4)	i	-0.01***	0.00
# Diagnosis for this patient (Below 90th percentile: < 18)	i	-0.04***	0.02***
# Diagnosis for this patient (Below 50th percentile: < 9)	i	-0.07***	0.05***
# Elderly in group (Above 10th percentile: > 150)	k	-0.04***	0.03***
# Elderly in group (Between 10th and 90th percentile: 150 to 30,000)	k	-0.05***	0.03***

Table 2.9: 2008 Robustness checks: Excluding outliers.

Specification changes:	Coefficient on elderly in diagnostic group	Coefficient on non-elderly in diagnostic group
Baseline	0.03***	-0.03***
Treatment v. non-treatment (as a test for extensive margin)	0.01***	-0.03***
Percent elderly instead of numbe elderly	0.01	n/a
ADJUSTMENT OF CONTROLS		
Excluding all i-level controls	0.03***	-0.03***
Excluding all k-level controls	0.03***	-0.02***
Ages 40 +	0.03***	-0.03***
All ages	0.02***	-0.02***
No hospital fixed effects	0.03***	-0.03***

Table 2.10: 1983 Robustness checks: Different specifications

the number of elderly within a patient's diagnosis, and a positive sign on the number of non-elderly. The 2008 data (Tables 2.8 and 2.9) and 1983 data (Tables 2.10 and 2.11) both remain true to the original form of the regressions in those years. The exclusion of fixed effects may not be strongly impacting the results because most hospitals are roughly in the middle of the pack in terms of treatment intensity. For example, using the three examples from Table 2.3, most hospitals have treatment effects most similar to Dominican Hospital. Table 2.4 shows a group of typical hospitals, which do not vary much from the treatment levels at Dominican. Indeed, no hospital in California comes close to the level of treatment at UCSF, and the rural hospitals in the sample are small enough not to have a huge impact either.

In addition to using the number of elderly as the key independent variable, it seemed worthwhile to try using the percent elderly. The main problem with this specification is that it precludes the possibility of including a control for popularity of an illness. Popularity has an ambiguous coefficient that depends on the percent elderly. The cutoff will depend on the particular relative reimbursement rates of Medicare and private insurers, information which is unavailable. Popularity is clearly an important factor in allocating investment dollars, so the regression remains somewhat lacking. This specification also does not match the theoretical model. Nonetheless, using percent elderly (Tables 2.8 and 2.10) had a negative coefficient in recent years and a positive coefficient in the early 1980's, consistent with the results when using number of elderly.

Excluding outliers:		Coefficient on elderly in diagnostic group	Coefficient on non-elderly in diagnostic group
No restrictions (Baseline)		0.03***	-0.03***
Avg # therapeutic procedures in group (10th to 90th percentile: 0.23 to 2.32)	k	0.04***	-0.03***
Avg # therapeutic procedures in group (Above 50th percentile: > 1.08)	k	0.11***	-0.07***
Avg # therapeutic procedures in group (Above 25th percentile: > 0.47)	k	0.05***	-0.05***
# Therapeutic procedures for this patient (Below 90th percentile: < 4)	i	0.03***	-0.03***
# Diagnosis for this patient (Below 90th percentile: < 18)	i	0.03***	-0.03***
# Diagnosis for this patient (Below 50th percentile: < 9)	i	0.03***	-0.03***
# Elderly in group (Above 10th percentile: > 150)	k	0.18***	-0.07***
# Elderly in group (Between 10th and 90th percentile: 150 to 30,000)	k	0.18***	-0.07***

Table 2.11: 1983 Robustness checks: Excluding outliers

2.8 Conclusion

Looking at a snapshot in time, Medicare's prospective payment system seems to preserve fair treatment of patients while forcing hospitals to consider costs. However, a full assessment must consider the dynamic impact of prospective payment on medical innovation. After all, innovation is what reshapes the face of health care in years to come. The results of this paper indicate that prospective payment is discouraging investment in therapeutic technology, since hospitals lose money for adopting new treatments for elderly patients' illnesses.

The empirical results reveal that patients with illnesses common among the elderly receive fewer therapeutic treatments today than patients with illnesses common among the young. This was not the case prior to the implementation of prospective payment in the early 1980's. In contrast, doctors treat Medicare patients equivalently or better than non-elderly patients in terms of treatment intensity. This supports the notion that hospitals exercise control over technology adoption while doctors make individual patient treatment decisions. The result is that incentives created through Medicare's payment system will show up in changes in R&D investment even if there is no difference in the way doctors treat the elderly.

These results have not been uncovered in previous analyses because previous studies have focused on shorter run treatment decisions rather than long run R&D incentives. Previous research on technology adoption in response to prospective payment has focused on different shares of elderly across hospitals. Differentiating between hospitals as a source of variation makes sense for short run or medium run analysis of technology adoption, but ignores the fact that in the long run, all hospitals interact with a single technological frontier. This paper's analysis utilizes variation across primary diagnoses in elderly share, assuming that each diagnosis has a unique technological frontier.

As new technology adds to the price tag of health care each year, policymakers may eventually want to realign incentives to channel medical R&D dollars toward the most cost-effective projects. President Obama already proposed creating an Institute for Comparative Effectiveness for health care technologies, though this was not part of the 2010 bill that Congress passed. Adjusting Medicare reimbursement rates to reflect cost-efficiency may be one potential tool for driving down the cost of Medical care through new technology. The results of this paper indicate that such a tool may be genuinely effective at re-channeling R&D dollars and reshaping the future of health care.

Chapter 3

Do Changes in Government Procedure Codes Impact Patient Treatment Patterns?

Abstract

This paper presents evidence that the addition of a new procedure code in the government's coding system leads to a sudden jump in the number of those procedures performed. It looks at a case study of a new procedure introduced in the 4th quarter of 2003, and looks at treatment patterns in years and quarters leading up to and after this change. The sudden change accounts for roughly a two percent increase in the probability that a patient in the relevant diagnosis groups will undergo a procedure. The magnitude of this change is about equal to about 90 additional procedures being performed per quarter in the state of California. This sudden increase in the probability of undergoing a procedure is most drastic for privately insured patients, Medicaid patients, and patients with managed care insurers. There is no sudden jump in the probability for Medicare patients, whose reimbursement depends only on diagnosis, independent of procedures. There is also no sudden jump among teaching hospitals, whose teaching responsibilities may outweigh reimbursement concerns.

3.1 Introduction

Both the government and private insurers use a government-issued coding system for hospitals to report diagnosis and treatment of patients. Every year, an independent board makes a few changes to this coding system. When a new procedure is added, it becomes easier for hospitals to bill insurance companies for that particular procedure.

Does the addition of a new code lead doctors to conduct more procedures on patients? This paper takes an empirical look by investigating one particular change that occurred

in late 2003. Before the change a “laparoscopic supracervical hysterectomy” (LSH) (Code: 68.31) would have to be coded under a more generic procedure code “subtotal abdominal hysterectomy” (Code: 68.3). After the change, the code 68.3 was retired, and subdivided into the two categories: the specific code already mentioned, and a generic code (68.39). Immediately after the change, about 22% of patients receiving either of the two new codes were categorized as receiving a “subtotal abdominal hysterectomy”.

The results of the study show that there was a sudden jump in a patient’s probability of receiving the relevant procedure in the quarter when the code was first introduced, 4th quarter of 2003. The project employs data that covers all patients discharged from California hospitals between 2001 and 2008. Over this time period, there was a constant rise in a patient’s probability of receiving a procedure in the umbrella category that includes the LSH procedure. Both before and after the new code was introduced, a patient’s probability of receiving such a treatment increased at about 0.2 percent per quarter. However, from 3rd to 4th quarter of 2003, the patient’s probability of receiving treatment increased a full 1.6 percent, from a 12.3 percent chance of treatment to a 13.9 percent chance of treatment. This jump was statistically significant.

For some groups of patients, the new code would make billing for the procedure much easier. For other groups, the new code would have little or no effect on reimbursement. Comparing the impact of the code change across these different patient types will help to tease out causality. For example, the existence of a new code would not impact reimbursement for Medicare patients, because Medicare’s reimbursement depends on the diagnosis code and is independent of the procedure code. Private insurers and MediCal, on the other hand, reimburse based on procedures performed, so a new procedure code makes it easier for hospitals to bill for their treatments. Patients paying out of pocket are not obligated to prove to an insurance company the validity of their procedure, so a new code would not likely impact their use of the procedure. Table 3.1 summarizes the anticipated effect of a new procedure code, depending on patient payer type, insurer management style, and hospital teaching status.

The results indicate that there was indeed a sudden jump in a patient’s probability of receiving the LHS procedure after the procedure code was introduced. This finding is robust to several variations on the model and to placebo tests more generally.

3.2 Background and Literature

Every year the government commissions an independent board to review the icd-9-cm (the International Classification of Diseases, 9th Revision, Clinical Modification) coding system and to make changes. Anyone can nominate a diagnosis or procedure to be added, but only a few changes will actually go into place each year. In addition, this board will also decide which codes to retire. The committee gathers information from doctors, hospitals and other stakeholders in making these decisions. Some new procedure codes represent procedures that

are new to the medical field and not yet adopted widely among doctors. Other nominations may include procedures that have become common practice, but have been previously billed under a broader umbrella category. The LSH procedure belongs to the later group.

No study has previously looked at changes in procedure codes under the icd-9-cm coding system. One study looked at the introduction of a new diagnostic code for the rotavirus in the icd-9-cm system (Parashar et.al, 1998). This study found that large hospitals and proprietary-owned hospitals were more likely to diagnose patients with relevant symptoms into the new diagnostic group. The authors suggested that doctors in smaller hospitals may have failed to adopt the new code at first and continued to use a more generic code for the rotavirus. Given this information, we might expect that the increase in usage of a procedure continues for a few quarters beyond the initial introduction of the code. The placebo test discussed in Section 5 indeed shows that placebos for a year and a half following the policy change are significant.

Some medical journals have featured articles that complain about the accuracy and usefulness of the icd-9-cm coding system. One complaint is that many of the codes do not come with specific clinical measures for proving that a patient belongs to the diagnosis group (e.g. Benesch et. al, 1997; McCarthy et. al., 2000; McIntyre et. al, 2008). This leads to a very uncertain judgement call on the doctor's part, and makes patient data imperfectly comparable across hospitals and doctors. However, most of these articles are concerned with diagnosis, not procedures. The main implication for this study is that there could be discrepancies in the patients belonging to the relevant diagnostic group for receiving an LHS procedure.

3.3 Methodology

3.3.1 Data

The data for this project came from California's Office of Statewide Health Planning and Development (OSHPD). It includes every inpatient discharge in the state of California from 2001 to 2008. This data set lists the quarter when the patient was admitted to the hospital, which allows the analysis to be disaggregated at the quarterly level to observe time trends and a discontinuity around the quarter when the new code went into place. For the analysis, I used only female patients between the ages of 35 and 45, since that is the demographic group who receive hysterectomies most commonly. I excluded patients over 45, because the rate of hysterectomies begins declining after that and would be more difficult to control for in a regression.

For the initial analysis, I include only patients who are likely candidates to undergo the LHS procedure. To identify this group, I looked at all patients receiving the new procedure, and looked at the top diagnosis associated with an LHS. About 77 percent of patients undergoing the LSH procedure had one of six diagnosis. Moreover, between the sixth and seventh most common diagnosis associated with LHS, there was a fairly large drop of in

patients. Therefore, I included in the sample any patient whose primary diagnosis was among the top six. Later on, when I conduct robustness checks, I look at a broader base of patients with any diagnosis, to check if there are any endogeneity problems in terms of selection into the diagnosis group.

Over time, the probability of treatment grew for this group of patients. In the first quarter of 2001, the relevant group of patients had a ten percent chance of undergoing a procedure in the general category of hysterectomy I am studying (code 68.3). By 2008, this percent had risen to about 18%. Based on the growth patterns observed after the new code for LSH was available, most of the growth in these procedures appears to be due to growth in LSH procedures, as opposed to other procedures under the same umbrella. Figure 3.1 shows these patterns over time.

3.3.2 Empirical Model

I employ a simple probit model to estimate a patient's probability of undergoing a procedure in the umbrella category for the LHS procedure. I could not directly estimate a patient's probability of receiving an LHS procedure, because no specific code for it existed prior to the 4th quarter of 2003 and the objective in the investigation is to determine any discontinuity in treatment that happened alongside this change. However, since a sizable share of patients (22 percent) in this umbrella category had LHS procedures, it is reasonable to believe that a discontinuity in the umbrella category is likely associated with LHS procedures. This assumption is bolstered by the fact that a patient's probability of undergoing a LHS procedure climbs steadily in the observable data, while other procedures under the same umbrella remain flat over the time period (Figure 3.1).

Equation 1 (Tables 3.2, 3.3, 3.4 and 3.5): probit model including only patients actually diagnosed:

$$proc = \beta_1 quarter + \beta_2 after + \beta_3 age + \varepsilon$$

where

proc is a patient's probability of undergoing the relevant procedure,

quarter is a time variable, and

after is a dummy for the time periods after the new procedure code was introduced.

As a second check on causality, I divided patients into several sub-groups that have different theoretical responses to the availability of a new procedure code. These predictions are summarized in Table 3.1 and are discussed with the results.

	Expected Effect	Reason
PAYER TYPE:		
Private insurer	Yes	Reimbursement may depend on procedure code
Medicare	No	Reimbursement does not depend on procedures performed
MediCal (Medicaid)	Yes	Reimbursement may depend on procedure code
Self-Pay	No	No insurer asking for code
INSURANCE		
MANAGEMENT STYLE:		
Managed Care	Yes	These insurers use procedure codes to screen for “appropriate care”
Traditional Payment	Depends	By definition, these insurers are less aggressive in screening for “appropriate care”, but they still may have some screening
HOSPITAL		
TEACHING STATUS:		
Not a teaching hospital	Yes	
Teaching hospital	Yes, but less than non-teaching	These hospitals have other goals and obligations in determining whether to conduct a particular procedure

Table 3.1: Predicted effect for different sub-groups

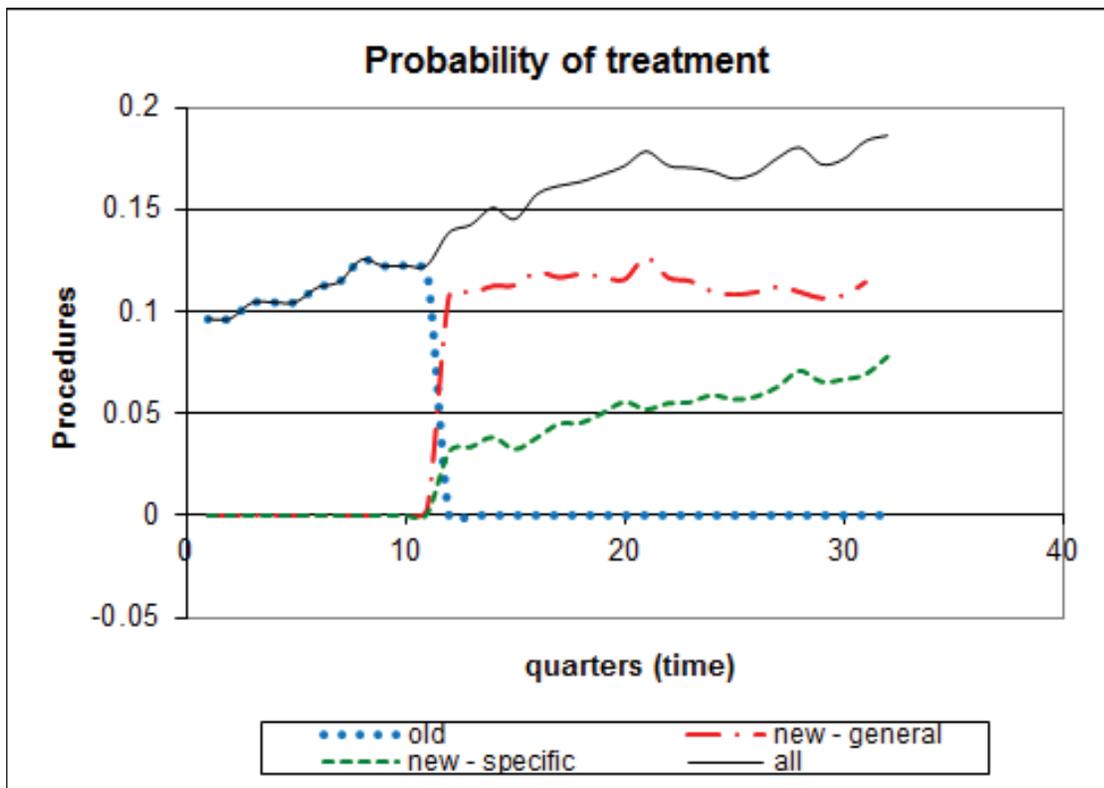


Figure 3.1: Probability that a patient undergoes a procedure
 The graph shows a patient's probability both before and after re-classification, given that the patient has one of six main diagnosis. There appears to be a small jump in the probability when the new procedure code is adopted.

3.4 Results

Tables 3.2 and 3.3 show that, controlling for time trends, there was a sudden jump in the probability of receiving a procedure after the new code became available. When looking at the raw data presented in Figure 3.1, we find that a sudden jump in the probability that a patient relevantly diagnosed will receive the procedure jumps from 12.3 percent to 13.9 percent in the quarter when the new code was introduced. This jump far outstrips the 0.2 percent average increase over all other time periods, and the jump can be detected with the naked eye in Figure 3.1. This is an economically important increase. About 30,000 women a year in California are diagnosed with one of the seven illnesses treated with the LSH procedure. This means that the new code caused an additional 90 procedures per quarter to be performed in California. The regression analysis verifies the statistical significance of this increase.

Tables 3.2, 3.3, 3.4 and 3.5 utilize a probit model and regress on whether or not the patient underwent a procedure. The after dummy is for time periods after the change in codes. The time variable is in quarters. Women ages 35 to 45 are included in these regressions, since they are the key group needing the procedure.

When running the regression on different sub-groups, the theoretical predictions outlined in Table 3.1 are supported by the empirical results. Most strikingly, the new code had no impact on the probability that a Medicare patient would undergo a procedure. This is what we would expect, since Medicare pays by diagnosis, not procedure. Private insurers and MediCal, who do bill by procedure, saw a sudden increase in the probability of undergoing a procedure immediately after the new code became available. Self-pay patients remained unaffected by the new code, as expected since they have no use for the code in billing an insurer.

Similarly, the increase in the probability of undergoing a procedure was bigger in magnitude for people insured via managed care companies. Managed care companies are more likely to refuse to pay for treatments, particularly if those treatments are difficult to document. The availability of the icd-9-cm code makes it easier to document, and also adds a certain stamp of approval to the performance of particular procedures. Insurers that reimburse via traditional payment are less likely to question bills submitted by the hospital, although that doesn't mean that there is no oversight at all. For these reasons, we would expect that the availability of a code would be more important for patients with managed care insurance than patients with fee-for-service insurance. Tables 3.4 and 3.5 indicate that the magnitude of the jump in probability was bigger for patients with managed care than for patients in general. These results show no significant jump for patients with traditional coverage, which is somewhat surprising and may indicate that the code's use in disputing claims and adding legitimacy is driving the results.

Finally, teaching hospitals have educational responsibilities that may be important in determining whether to perform a procedure, even if it is difficult to collect from insurers. The expectation, then, was that the effect would be smaller in magnitude for teaching

	Baseline	Privately insured	Medicare	MediCal (Medicaid)	Self-Pay
Time variable (ln)	0.070***	0.077***	-0.050	0.034	0.122
After dummy	0.117***	0.118***	0.274	0.159*	-0.041
Constant	-4.564***	-5.345***	-3.247	-0.487	-2.178
N	45,143	35,445	784	6,024	635

Table 3.2: Using patients actually diagnosed

	Baseline	Privately insured	Medicare	MediCal (Medicaid)	Self-Pay
Time variable (ln)	0.037*	0.070***	-0.014	-0.040	-0.159
After dummy	0.130***	0.139***	0.364	0.214**	0.030
Constant	1.799***	-2.853***	12.052***	6.199***	7.060
N	34,912	21,266	2,148	8,451	954

Table 3.3: Using patients among the top 4% most likely to be diagnosed (predicted)

	Baseline	Managed Care	Traditional Coverage	Teaching	Non-teaching
Time variable (ln)	0.070***	0.071***	0.092**	0.012	0.080***
After dummy	0.117***	0.129***	0.064	-0.113	0.139***
Constant	-4.564***	-5.161***	-1.697*	-8.360***	-4.215***
N	45,143	36,312	7,769	4,513	39,700

Table 3.4: Using patients actually diagnosed

	Baseline	Managed Care	Traditional Coverage	Teaching	Non-teaching
Time variable (ln)	0.037*	0.052**	0.064	0.025	0.039
After dummy	0.130***	0.155***	0.050	-0.069	0.157***
Constant	1.799***	-1.615***	7.080***	5.944***	0.939*
N	34,912	23,198	10,395	4,180	29,373

Table 3.5: sing patients among the top 4% most likely to be diagnosed (predicted)

hospitals. Still, you would expect some effect for teaching hospitals. The data shows no significant effect for teaching hospitals. This is somewhat surprising and may indicate that teaching responsibilities at these hospitals overshadow reimbursement concerns in treatment decisions.

3.5 Robustness Checks

One concern with the model described above is that there may be endogeneity if there is a sudden change in selection into the diagnosis group. It is possible, that prior to the introduction of the new procedure code, doctors favored certain diagnosis as a way of signaling that they were performing the LHS procedure. If this is the case, then after the code became available, they may have felt more free to assign other diagnosis as the primary diagnosis. The result would be a sudden shrinking in the pool of patients in the diagnosis group that the analysis looks at. Such a shrinking would lead to an inflated jump in the probability of treatment that was caused by a change in who is diagnosed rather than who undergoes the LHS. This is not necessarily likely, but the fact that it is possible makes it worthwhile to check into alternative ways of modeling.

To address this, I used data prior to the policy change to predict a patient's probability of selecting into the diagnosis group. This prediction was based on the patient's secondary diagnoses, age, race and ethnicity. The correlation coefficient between a patient's predicted probability of being diagnosed and their actual diagnosis is 42%. After coming up with predictions, I ran the original model, including the group of patients most likely to be diagnosed based on the prediction. I used 8% probability of diagnosis as the cutoff, which accounts for the top 5% of patients most likely to be diagnosed were in that group. This cutoff was chosen because it was the closest to the actual number of patients diagnosed.

Equation 2: Probit. Using only data prior to the change in code. Creating a prediction of whether or not a patient will be among the diagnosis group.

$$diag = \beta_0 + \beta_1 quarter + \beta_2 race + \beta_3 ethnicity + \beta_4 SecondaryDiag + \varepsilon$$

where

diag is a dummy for whether or not the patient was diagnosed with one of the six most common diagnosis associated with the procedure, and

SecondaryDiag is a vector of dummies for whether or not the patient received any of the top ten secondary diagnosis associated with the diagnosis commonly associated with the procedure.

Tables 3.3 and 3.5 show results from the same regression indicated in Equation 1, except instead of including patients actually diagnosed, it includes all patients with a probability of being diagnosed that is above the cutoff. In both tables, the magnitude and significance of the coefficient on the dummy for after the code change was introduced greatly resembles

the original regression. This indicates that there wasn't a sudden drop in the probability of being diagnosed that is driving the sudden increase in a diagnosed patient's probability of being treated. A sudden increase is persistent for time periods after the change in codes. The time variable is in quarters. Women ages 35 to 45 are included. A patient's probability of being diagnosed, which might be the expected effect after the new code became available, would bias the results in the other direction, strengthening the argument that there was a sudden bias in favor of treatment.

It should also be noted that another possibility would be a sudden increase, rather than a sudden decrease, in patients diagnosed in the seven illnesses associated with the LHS. This would happen if doctors wanted to diagnose someone in order to justify performing a procedure that now has a higher reimbursement. If this were the case, it would not weaken the validity of the primary analysis. Rather a sudden increase in the group of people diagnosed that is also associated with a jump in the probability each is diagnosed means that there are definitely more procedures being performed due to the new code.

Another concern might be the timing of the change. Maybe doctors anticipated the change and therefore put off the surgery for a month or two. This will be important to test for, because Figure 3.1 will show the probability of undergoing a procedure is more or less flat in the couple of quarters leading up to the change. To make sure that the result isn't merely driven by a change in timing, I ran the regressions excluding the two quarters before and the two quarters after the new code was introduced. This did not change the results at all. There was still a significant jump of a very similar magnitude.

To test robustness, I ran a placebo test using each quarter leading up to and following the policy change. Figures 3.2 and 3.3 graph these coefficients with 95 percent confidence intervals, and shows that the test would have yielded significant results with placebos in some of the quarters after the policy change, but not generally in the quarters before the policy change. The quarter immediately before the policy change would have been significant as a placebo. However, this may be due to timing. In particular, to disaggregate the data to the quarterly level, I had to use the patient's date of admission. Some patients may have been admitted in the quarter prior to the policy change but treated in the quarter after the policy change. These patients would have shown up in the quarter prior to the policy change and may be driving significance in that placebo test.

Figures 3.2 and 3.3 show the results from running placebo tests. These graphs plot coefficient on the placebo for time periods (quarters) other than the quarter of change. The red data point represents the coefficient in the quarter where the new code was introduced. The light lines show 95% confidence intervals.

The fact that the placebos in the year and a half following 4th quarter 2003 are still significant is somewhat concerning. However, this may be due to the fact that information about the new coding change may have taken a while to diffuse to doctors. For example, doctors are not the ones to enter the new codes. Instead hospital administrative staff will be the ones to conduct the billing. The hospital administration may have taken time to realize that these procedures were more lucrative and easier to collect reimbursement. It may have

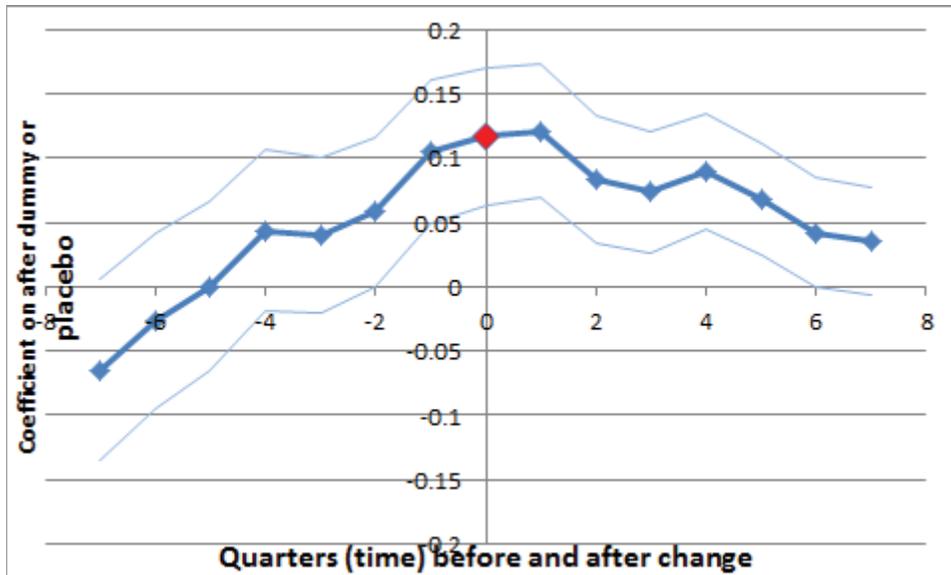


Figure 3.2: Using patients actually diagnosed

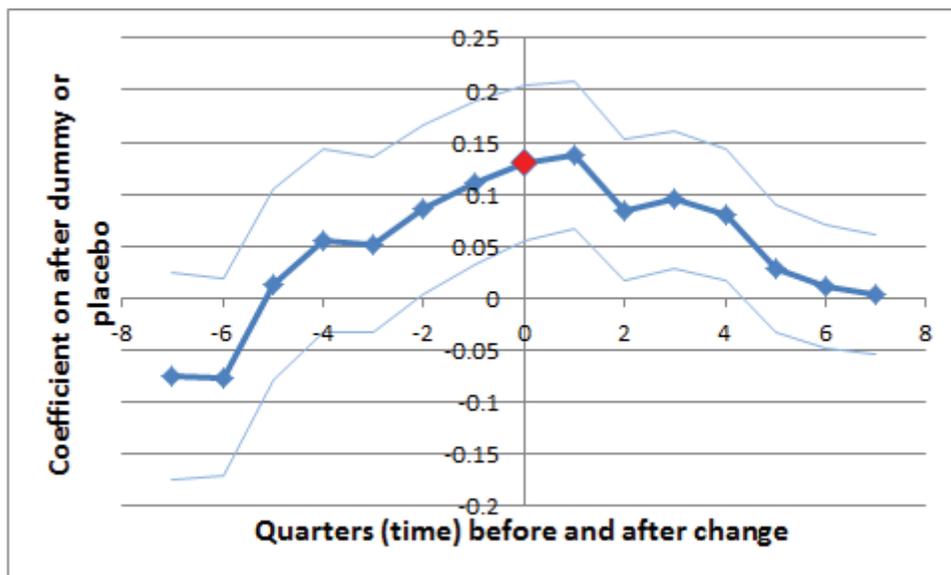


Figure 3.3: Using patients among the top 4% most likely to be diagnosed

taken additional time for the hospitals to convince doctors to change their behavior as a result, particularly if the means of influencing doctors was indirect.

Finally, I considered the possibility that the availability of a code may have increased the pace at which the procedure was adopted by doctors. Accordingly, I included an interaction variable between time and the after dummy. This had no impact on the coefficient size or significance for the original after variable. Also, the interaction term was insignificant.

3.6 Conclusion

The results of this study indicate that when the government introduced new procedure code for the “laparoscopic supracervical hysterectomy”, hospitals and doctors perform more of that kind of procedure. The magnitude of this change accounted for about 90 additional procedures per quarter being performed in California after the code was introduced. Whether this increase represented a positive or negative change for women is beyond the scope of an economist’s reach. Only doctors can determine the right level of treatment. This paper simply concludes that the availability of procedure codes is impacting hospitals’ rate of treatment. This fact becomes more evident in light of the patient group breakdowns I looked at. When breaking patients down into subgroups that would be likely and unlikely to be impacted by the new code, we find the expected jump only among groups whose insurance reimbursement depends on icd-9-cm procedure codes.

This project does not distinguish the reason for this increase, and several possibilities are at play. Having a code available makes it easier for hospitals to bill insurers. It also makes it more difficult for insurers to question the validity of a procedure or dispute a claim for legitimacy reasons. In some ways, adding a procedure to the icd-9-cm code list makes it part of the medical cannon in a way that is literally approved by the government.

Of course, the jump in treatments could be due to a change in payment incentives associated with the new procedure. Undoubtedly insurers would have to assign a reimbursement rate for the new procedure. If that particular insurance company did not previously have a standard rate of reimbursement for the LSH procedure, there would have been a reimbursement change. If this change was, on average, higher than previous reimbursements, we would expect to see such a jump even if the ease of billing were not a factor. We do not have specific information about insurer reimbursement rates, and therefore can only surmise about payment incentives.

There are several arenas where this could be relevant from a policy standpoint. First, the committee that determines new procedure codes will want to consider the impact of a sudden increase in a procedure after they introduce a new code. If it is a procedure that is gaining popularity without medical research backing its effectiveness or safety, the committee may want to wait for a few years before introducing the code and propelling forward the diffusion of the new procedure.

Second, this information can be used to better predict costs and procedures going forward.

States running Medicaid programs and private insurance companies may want to take into account new procedure codes when running cost projection analysis.

Finally, introducing new procedure codes could potentially be a tool for influencing diffusion of new procedures. Specifically, the timing of the introduction of a new code could impact how quickly that procedure gets adopted by new doctors and hospitals. Procedures with proven safety, effectiveness and cost-effectiveness could be given a boost by hastening diffusion through the introduction of a new code early in the game. If drawbacks to a procedure become apparent early on, the review committee can hold off on introducing the code until a broader consensus has been reached in the scientific community.

Future research will need to look into other code changes to see if similar increases in treatment are associated with them. It also may be relevant to study how the magnitude of an increase in treatments correlates with the extent to which a procedure has already diffused in usage. For example, it could be the case that there is a bigger jump for procedures that are already well established in medical practices, or else a bigger jump for more newly introduced procedures. Understanding the effect of code changes on the spread different types of procedures could be key in eventually using it as a policymaking tool.

Chapter 4

Health Care Supply and County-Level Variation in ADHD Prescription Medications

Joint with Tim Bruckner, Chris Brown Mahoney, Brent Fulton, Peter Levine, and Richard Scheffler

Abstract

Purpose: Although much literature reports small-area variation in medication prescriptions used to treat Attention-Deficit Hyperactivity Disorder (ADHD), scant research has examined factors that may drive this variation. We examine, across counties in the U.S., whether the use of prescription medications to treat ADHD varies positively with supply-side health care characteristics. **Abstract:**

Methods: We retrieved annual prescription data for ADHD medications in 2,734 U.S. counties from a nationally representative sample of 35,000 pharmacies in 2001 to 2003. We used a county-level, multivariable fixed effects analysis to estimate the relation between annual changes in health care supply and ADHD medication prescriptions. Methods controlled for time-invariant factors unique to each county as well as ADHD prevalence.

Results: From 2001 to 2003, retail prescription purchases for ADHD medications increased 33.2%. In the multivariable analysis, ADHD medication prescriptions move positively with an increase in the concentration of total physicians. In addition, ADHD medication prescriptions move inversely with changes in the percent non-Hispanic black population.

Conclusions: Supply-side health care factors may contribute to the rise from 2001 to 2003 in ADHD medication prescriptions. This finding warrants concern because it implies that the relative capacity of the health care system affects population prescription rates.

We encourage further exploration of the contribution of the supply-side of the health care market to secular changes in ADHD medication prescriptions.

4.1 Introduction

Attention deficit hyperactivity disorder (ADHD) represents the most commonly diagnosed childhood behavioral disorder in the U.S (Williams, et. al., 2004). This psychiatric condition predisposes children to reduced academic achievement (Barkley et. al., 1990; Hindshaw, 2002), suboptimal social development (Bagwell et. al., 2001), and increased risk of accidental injury (Rowe, Maughan, and Goodman, 2004). The persistence of these impairments into adolescence and adulthood, combined with the over 400% increase in the 1990s of the use of medication to treat ADHD (Thomas et. al., 1994; Bhatara et. al., 2004), has engendered much public health concern (Eberstadt, 1999; Mayes, Bagwell, Erkulwater, 2008).

The child's environment exerts a strong influence on the likelihood of ADHD diagnosis. Environmental factors include parental and school characteristics, health insurance, cultural tolerance, and public policies (Wasserman et. al., 1999; Bussing, Schoenberg, Perwien, 1998; Bussing et. al., 1998; Schneiders et. al., 2003; Ford, Goodman, Meltzer, 2004). Variability of these contextual factors reportedly contributes to large differences in ADHD prevalence and treatment across geographic region. For example, although the most recent national ADHD prevalence estimates range from 3 to 8 percent (Mental health in the United States, 2003), individual community studies have reported from 1.7 to 26 percent prevalence (Rappley, 2005; Zuvekas, Vitiello, Norquist, et. al., 2006). Also consistent with this large variability, consumption rates for 24 medications used to treat ADHD vary as high as 1:10 between communities within states (Spanos, 1996; Wennberg J, Wennberg D, 2001).

The literature on geographic variation in the U.S. reports a positive association between supply-side characteristics of the health care system (e.g., physicians per capita) and health care demand (Wennberg, 2002). Scant work, however, has examined whether health care supply-side factors play a role in the small-area variation in ADHD medication prescriptions (Rappley et. al., 1995). Patients must find a prescribing physician in order to receive psychostimulants. A higher concentration of prescribing physicians in a county may improve access to medical care. This circumstance could increase the number of ADHD medication prescriptions. If supply-side factors contributed to geographic variation in ADHD prescriptions, we would expect a positive relation between prescriptions sold and indicators of access to physicians, (e.g., per capita concentration of total physicians, percent of pediatricians, and percent child psychiatrists in a county). To our knowledge, no research has examined this supply-side issue in an analytic setting that controls for unobserved regional factors or accounts for the often-reported influence of sociodemographic, policy, and school-level factors on ADHD medication sales.

Quantifying the relation between health care supply and prescriptions to treat ADHD represents an important policy issue for two reasons. First, an increase in the use of costly

prescription medications continues to strain the national health care system. Adderall and Strattera, two long-acting ADHD medications made available on the market within the last decade, exemplify a key source of these rising costs (Scheffler et. al., 2007). Identification of system-level health care factors that drive ADHD medication sales may help contain health care costs (Barkley, 2003).

Second, the literature does not converge on whether ADHD medication use appears equitable across racial/ethnic and socioeconomic groups. For instance, psychotropic medication use among non-Hispanic black children reportedly falls below the level of non-Hispanic white children (Stevens, Harman, Kelleher, 2005). However, analyses in Canada find that children in low-income areas appear twice as likely as children in high-income areas to receive ADHD medication (Brownell, Mayer, Chateau, 2006). Examination of key demographic variables in a multivariable setting may uncover whether the assertion of inequitable ADHD medication use holds at the population level across diverse regions with different health care characteristics.

We set out to test, across all counties in the U.S., whether health care supply-side factors vary positively with ADHD medication use. We advance earlier work in three important ways (Bussing, Schoenberg, Perwien, 1998; Fulton et. al., 2009; Bokhari, Mayes, Scheffler, 2005; Cox et. al., 2003; Cohen, Hesselbart, 1993). First, we capitalize on a unique, nationally representative dataset of pharmacy prescriptions for ADHD medications for U.S. counties from 2001 to 2003. Second, we examine concurrently key health care supply-side, sociodemographic, school, and ADHD policy variables. Third, our longitudinal data permit an analytic strategy which removes time-invariant regional confounders that could drive ADHD medication prescriptions.

4.2 Methods

Our analysis required retrieval of data from multiple sources. Table 4.1 lists the source file for each variable. We acquired data on medications sold to treat ADHD for the years 2001, 2002, and 2003 from Wolters Kluwer (formerly NDCHealth). The data include information at the zip code level on 24 medications used to treat ADHD. Wolters Kluwer uses a weighted, stratified sampling strategy of 35,000 pharmacies in all U.S. zip codes to estimate regional retail prescription sales and tablet quantity of medications sold. These estimates reflect ADHD medication sales from 97% of all retail pharmacies in the U.S., including independent, chain, supermarket, and specialized pharmacies.

To our knowledge, the Wolters Kluwer data comprise the most valid estimate of prescriptions sold for ADHD medications with geographic resolution smaller than at the state level. The consistent methodology of data collection permits valid comparisons across places and times. The Wolters Kluwer data contain information on the quantity of capsules purchased for the ADHD medications. The dosage of these capsules depends on the medication brand name and type. For instance, a one milligram dose of Adderall, which comprises an 8.9

Variable	Source	Coverage
Standardized Mg of Ritalin	Wolters Kluwer	2001-2003
ADHD prevalence	CDC's National Survey of Children's Health	2003
Physicians by specialty and age; public insurance coverage	Area Resource File	2001-2003
Health Maintenance Organization (HMO) penetration	Health Leaders InterStudy Publications on HMOs	2001-2003
Race/Ethnicity, Per Capita Income, Age distribution, population	U.S. Census	2001-2003
Student to teacher ratio, Individualized Education Programs	Dept. of Education's Common Core of Data	2001-2003
State school accountability measures	Education Week	2001-2003
State ADHD law against teacher referral	ablechild.org	2001-2003

Table 4.1: List of Data Sources

Variables used in the Analysis. For the analysis, we applied 2003 prevalence estimates to earlier years of county data (2001 and 2002).

percent share of the ADHD market from 2001 to 2003, equates to a 3.29 milligram dose of Ritalin. This equivalence implies that a 1 milligram dose of Adderall lasts 3.29 times as long as a 1 milligram dose of Ritalin. To allow comparability across counties of total quantity purchased, an expert pediatrician on our research team (Levine) standardized the dosages of all ADHD medications to milligrams of Ritalin. Table 4.2 shows the milligram equivalent of Ritalin for each 1 milligram of the 24 medications. We used these standardized milligrams of Ritalin for all analyses.

We aggregated the five-digit zip code data up to the county ($n=3,140$) to permit analyses at the county level. Prior analyses of small area variation find that counties comprise a more geographically relevant unit as compared with 3- or 5- digit zip codes (Wennberg J, Wennberg, 2001). For zip codes that span across multiple counties, we partitioned the data into multiple counties using census weights that reflected the proportion of the population in the zip code that resides in each county (US Census Bureau, 2008). Next, we excluded counties with missing annual values of medication sales or independent variables ($n= 406$), which yielded data on 2,734 counties.

4.2.1 ADHD Prevalence by County

Diagnosed ADHD prevalence varies by geographic region, and prevalence appears correlated with ADHD medication use (Schneider, Eisenberg, 2006). We know of no survey that estimates ADHD prevalence by county. As a surrogate, to control for prevalence we used the 2003 Centers for Disease Control and Prevention's National Survey of Children's Health, the largest and most recently available survey on children.¹⁶ This nationally representative survey provides, for each state, gender and race-specific estimates of ADHD prevalence for children aged 5 to 17 years. We used indirect standardization methods, recommended in the epidemiologic literature, and multiplied for each state the National Survey of Children's Health's race- and gender-specific prevalence measures by each county's race- and gender-specific counts of children (Kleinbaum, Kupper, Morgenstern, 1982). Summation of these products yielded the estimated number of children in each county with an ADHD diagnosis. We used this county-level estimate as the denominator in the dependent variable for all three years of data, and assumed a stable ADHD prevalence from 2001 to 2003.

4.2.2 Dependent Variable

We used as the dependent variable the ratio of standardized milligrams of Ritalin sold per year to the estimated number of diagnosed children with ADHD. We used ADHD prevalence estimates by county for 2003 as the denominator of the dependent variable for all three years of data (e.g., 2001-2003). The dependent variable gauges the annual dosage per child diagnosed with ADHD.

1 mg =	mg Ritalin
Adderall tablet	2.86
Adderall Extended Release	2.14
Concerta	0.69
Cylert	0.44
Desoxyn	1.75
Dexedrine spansule	2.14
Dexedrine tablet	1.75
Dextroamphetamine tablets	1.75
Dextroamphetamine spansule	2.14
Dextrostat	1.75
Focalin	2.00
Metadate-controlled delivery	1.25
Metadate-Extended Release	0.83
Methylin Extended Release	0.83
Methylin tablet	1.00
Methylphenidate tablets	1.00
Methylphenidate-Slow Release	0.83
Mixed Amphetamine Salt Tab	2.86
Pemoline	0.44
Provigil	0.28
Ritalin	1.00
Ritalin-Long Acting	1.25
Ritalin-Slow Release	0.83
Strattera	0.83

Table 4.2: ADHD medication prescriptions
Prescriptions used in the NDC dataset and their 1 milligram equivalent in milligrams of Ritalin

4.2.3 Health Care

We retrieved key independent variables from the area resource file (ARF) and the Health Leaders InterStudy Publications on HMOs (US Department of Health and Human Services, 2008; InterStudy Publications, 2008). The ARF contains data on health professions, whereas the InterStudy Publication on HMOs provides annual estimates of the number of operating HMOs and enrollment figures. We selected the physician and health care supply-side factors previously reported in this journal to vary with ADHD medication use (Bokhari, Mayes, Scheffler, 2005). We, consistent with this earlier work, included total physicians per capita, percentage of physicians that are pediatricians, percentage of physicians that are child psychiatrists, the percentage of physicians less than 45 years of age, and HMO concentration (i.e., the per capita number of HMOs). We also retrieved information on the percent of the population with public health insurance.

Based on the descriptive findings of Bokhari and colleagues, 28 we hypothesize a positive relation between ADHD medication prescriptions sold and per capita concentration of total physicians, percent of pediatricians, child psychiatrists, young physicians, and HMO concentration. These variables may gauge the extent to which counties have improved access to medical care, which in turn could increase the number of ADHD medication prescriptions. We hypothesize an inverse relation between such prescriptions and percent of the population with public insurance.

4.2.4 Sociodemographic Variables

The literature reports substantial variation in ADHD medication prescriptions across race, age, and income level (Bussing et. al., 1998; Schneiders et. al., 2003; Ford, Goodman, Meltzer, 2004; Zuvekas, Vitiello, Norquist, 2006). We used the U.S. Census county-level estimates of these sociodemographic characteristics for the three years under study. As described above, we adjusted for the race, age, and gender of children in each county. We also included as covariates mean per capita income, percent non-Hispanic black, and percent Hispanic population.

4.2.5 School Variables

We acquired county-level school data from the Department of Education's Common Core of Data, which provides a comprehensive national statistical database of all public elementary and secondary schools and school districts (National Center for Educational Statistics Common Core of Data, 2008). The consistent and standardized method of data collection permits comparisons across states and years. We specified in the analysis the student to teacher ratio, as well as the proportion of children in an individualized education program. Individualized education programs serve children with learning or other disabilities and may gauge the school's capacity to provide counseling to children with ADHD. Next, we gath-

ered information on four state school accountability measures that reportedly vary with receiving an ADHD diagnosis (Schneider, Eisenberg, 2006): state rewards high-performing or improved schools, state sanctions low-performing schools, state assigns ratings to schools, and state has school report (Education Week, 2008). State school accountability measures are structured to reward highperforming and sanction low-performing schools on the basis of academic achievement. The presence of such measures may compel teachers and schools to recommend an ADHD diagnosis that could lead to effective treatments, which may improve academic performance and reduce problem behaviors in that school. Also at the state level, we specified in the analysis laws primarily designed to prohibit school personnel from recommending that a child take a medication for ADHD.

4.2.6 Data Analytic Procedures

We conducted a multivariable, fixed effects analysis at the county-level. This approach arrives at annual change values across the time frame (e.g., 2001 to 2002, and 2002 to 2003) and examines whether annual changes in ADHD medication rates move with annual changes in the independent variables. Omitted county-level variables that are relatively stable over time may bias effect estimates if they correlate with supply-sensitive health care characteristics and influence ADHD medication prescriptions. These variables could include, for example, cultural norms, community resources, patient preferences, or medication price. To control for this potential bias, we included county fixed effects. This approach permits estimates of the effect of a change in explanatory variables on a change in the dependent variable. We also specified year effects to control for generally occurring time trends in ADHD medication prescription patterns. We used the “panel-corrected” standard error option to allow for nonindependence of repeated observations, contemporaneous correlation, and efficient estimation in the presence of heteroskedasticity.

4.3 Results

From 2001 to 2003 individuals purchased approximately 85 billion standardized milligrams (mg) of ADHD medication from retail pharmacies. Retail prescription purchases increased 33.2% from 2001 to 2003. County-level pharmacy prescriptions for ADHD medications varied substantially (Table 4.3). The mean level of ADHD medication consumption per diagnosed child (i.e., 4,340 mg) appears consistent with estimates based on dosing recommendations and adherence rates. Health care supply characteristics also exhibit small-area variation; for example, the inner-quartile range of physicians per 1,000 persons spans from 0.5 to 1.6. The far right columns of Table 4.3 display the mean, standard deviation, and inner quartile range of the annual change values—the within-county variation that we exploit in our fixed-effects analysis. We observe a mean annual variation of 16 percent (i.e., 695 / 4,340) in the dependent variable. The health care supply variables exhibit annual variation that typically

ranges from one to ten percent.

Inspection of the panel error structure of our data indicated heteroskedasticity and serially correlated errors (at lag 1year). We, therefore, specified an estimation strategy that efficiently estimates standard errors in the presence of these conditions. Table 4.4 shows the regression results. Estimated ADHD prevalence per county is the denominator of the outcome variable. We remind the reader that differences in the metrics used for the independent variables (e.g., student to teacher ratio, concentration of physicians) precludes a direct comparison of the relative strengths of coefficient estimates across all variables. The year indicator variables for 2002 and 2003 show a positive relation with ADHD prescriptions, suggesting a strong upward time-trend. Changes in the percent non-Hispanic black population move inversely with changes in ADHD medication prescriptions. County per capita income moves positively with medication rates. In addition, one school characteristic — the concentration of Individualized Education Programs—is associated with fewer ADHD medication sales.

Consistent with the notion of supply-side geographic variation, health care supply predicts county-level ADHD prescriptions (Wennberg, 2002). An increase in total physician concentration, as well as child psychiatrists, moves with annual increases in ADHD prescriptions. HMO penetration, however, moves inversely with county prescriptions.

To give the reader an estimate of the magnitude of our discovered health care findings, we calculated the effect on change in ADHD medications statistically attributable to changes in physician supply. A modest increase, *ceteris paribus*, in the county concentration of physicians—from 1.2 per 1,000 persons (i.e., the mean level) to 1.4 per 1,000 persons—implies a 340 standardized mg Ritalin increase in prescriptions per child with ADHD. This equates to about an 8% increase above the mean level (i.e., 4,340 mg; Table 4.3) of medication purchased. An alternative interpretation is that a modest increase in physicians may increase diagnosed ADHD prevalence while holding steady the number of prescriptions per child.

We assessed the robustness of our results by removing outliers in ADHD medication prescriptions. We defined outlier counties as those with changes over time in ADHD medication prescriptions below the 1st percentile or above the 99th percentile of the distribution of change values in county prescriptions. Statistical inference for all but one of the coefficients in Table remains the same as in the original test (child psychiatrists as a percentage of all physicians became nonsignificant). Results also do not appear sensitive to inclusion of other demographic control variables (e.g., percent of population below poverty line).

4.4 Discussion

Nationally representative data from retail pharmacies in 2,734 counties indicate that prescriptions for medications to treat ADHD vary positively with changes in the concentration of physicians. Findings remain robust to control for other county-level health care variables, ADHD prevalence, racial/ethnic composition, mean per capita income, and school characteristics. Our results show that an increase in the concentration of health care professionals

Variable	Mean	SD	IQR	Mean annual change	SD of change	IQR of change
Standardized Mg of Ritalin sold per ADHD child	4,340	5,480	622 - 6,163	695.0	1658.1	8.9 - 1,113.6
Physicians per 1,000 persons	1.2	1.4	0.5 - 1.6	0.024	0.13	-0.028 - 0.075
Pediatricians (% of MDs)	4.8	6.8	0.0 - 7.3	0.019	3.2	-.17 - 0.11
Child psychiatrists (% of MDs)	0.41	1.3	0.0 - 0.35	-0.0029	0.68	0.0 - 0.0
% of MDs < 45 years	32.6	17.7	23.2 - 41.6	-1.00	9.1	-3.0 - 1.2
HMOs per 1,000 persons	0.13	0.17	0.03 - 0.16	-0.013	0.086	-0.027 - 0.0038
% population with public insurance	30.8	9.4	24.6 - 36.4	0.063	4.5	-1.8 - 1.9
Student teacher ratio in public schools	14.8	6.2	13.1 - 16.3	-0.24	7.1	-0.5 - 0.2
Total Individualized Education Programs per 1,000 persons	25.1	6.7	20.8 - 28.4	0.19	2.3	-0.59 - 1.0
State school accountability measures (range: 0 to 4)	2.38	1.27	1 - 4	0.05	0.85	0 - 0
Per Capita Income (\$)	23,972	6,357	20,548 - 26,948	778.4	1288.3	81.5 - 1336.5
% population black	8.8	14.3	0.4 - 10.2	0.021	0.21	-0.0043 - 0.065

Table 4.3: Descriptive county characteristics, 2001 to 2003. Mean, standard deviation, and inner-quartile range for both annual values and “change” values from year-to-year. (Year-to-year change values reflect the difference from 2001 to 2002, and 2002 to 2003. We use these annual change values for the statistical analysis)

	Coef.	(SE)
Supply Side Characteristics		
Physicians per 1,000 persons	1,699.32	(274.48)***
Pediatricians as a % of all physicians	17.22	(10.55)
Child psychiatrists as a % of all physicians	143.26	(61.12)*
% of physicians < 45 years of age	- 5.42	(3.37)
Insurance Characteristics		
HMOs per 1,000 persons	-1673.13	(751.17)*
% of population with public insurance	-8.16	(7.85)
School Characteristics		
Student to teacher ratio in public schools	-1.70	(20.41)
Total Individualized Education Programs per 1,000 persons	-49.36	(17.35)**
State school accountability score	-71.48	(49.70)
State ADHD law against teacher referral	-183.42	(135.19)
Population Characteristics		
Per Capita Income (in \$1,000s)	59.41	(27.54)*
% non-Hispanic black population	-1652.29	(251.12)***
% Hispanic population	- 917.43	(750.27)
Year Effects (2001 as referent)		
2002	303.49	(58.93)***
2003	1250.36	(59.96)***

Table 4.4: Fixed effects regression results

Fixed effects regression results predicting a change in standardized milligrams of medication sold per ADHD child per year in U.S. counties from 2001 to 2003 (panel-corrected standard errors in parentheses). Note: All tests of significance are two-tailed. * $p < .05$ ** $p < .01$ *** $p < .001$. A standardized milligram of medication is equivalent to one milligram of Ritalin. County fixed effect parameter estimates are not displayed.

predicts an increase in ADHD medication rates.

Findings appear consistent with the literature which documents substantial (i.e., 10 to 1) small-area variation in ADHD medication prescriptions (Wennberg J, Wennberg D, 2001; Bokhari, Mayes, Scheffler, 2005). Unlike previous work, however, our analytic approach sheds light on potential explanations of this variation. We used panel data over a three year time span, employed a fixed-effects approach, and adjusted for school, policy, and sociodemographic variables. We minimize bias due to relatively time-invariant omitted variables (e.g, physician practice styles, cultural norms) which typically are not controlled in cross-sectional studies and may account for geographic variation in prescription patterns.

The discovered phenomenon of supply-side care for ADHD medication prescriptions warrants attention because it implies that the relative capacity of the health care system influences medication rates. Wennberg (Wennberg, 2002) and others (Fisher et. al., 2000) note that clinicians generally try to fill appointments to full capacity. As this relates to ADHD, the increase of physicians per capita in a county may reduce the interval between a child's visits and/or increase the likelihood of ADHD diagnosis. These factors could increase medication use even if the true underlying prevalence of ADHD across counties does not differ appreciably. Increasing the concentration of physicians, therefore, may affect diagnostic patterns as well as prescription rates.

The inverse relation between HMO penetration and ADHD medication rates diverges from an earlier descriptive report by Bokhari and colleagues (Bokhari, Mayes, Scheffler, 2005). One explanation for this finding involves the HMOs' objective to minimize health care costs. HMOs may promote a provider culture that limits prescriptions with relatively expensive formulations. This circumstance may lower medication consumption levels in areas with an increasing concentration of HMOs.

Our race/ethnic findings indicate inequitable use of ADHD medications among non-Hispanic black, relative to non-Hispanic white, populations. This disparity may arise due to reduced diagnostic prevalence, reduced medication seeking once diagnosed, or both. Policies that endeavor to optimize use of medications to treat ADHD may want to target counties with non-Hispanic black populations.

The strong upward trend from 2001 to 2003 in per capita ADHD medication prescriptions converges with national reports of increased ADHD prescription rates among children since 2002. Explanations for this increase include that technological innovation of new formulations (e.g., Adderall Extended Release) may have promoted adherence among children already diagnosed with ADHD. Specifically, the availability of once-a-day formulations, as opposed to three-a-day formulations, may improve adherence. Once-a-day formulations may reduce the social stigma of taking medication since children may no longer ingest the medication during school hours.

Limitations of small area analyses include that findings may not generalize to individuals. We caution against using our ecological coefficients to estimate individual ADHD medication prescriptions. The reader should consider our results as contributing to the policy debate on developing uniform standards for diagnosis and treatment of ADHD across regions and

cultures, rather than informing clinical interventions or elucidating individual behavior. In addition, lack of county-specific data on ADHD prevalence for 2001 and 2002 led us to assume constant prevalence across the test period. This circumstance implies that a secular rise in county-level ADHD prevalence that coincides with increasing county-level physician concentration may contribute to our discovered supplyside effects. We also do not have county data on other variables known to affect ADHD medication use (e.g., mental comorbidities or parental or family support) or on what fraction of the purchased medications children actually consumed. In addition, we cannot rule out the rival explanation that unmeasured factors that trend in a county, but are not caused by health care supply variables, account for the findings. Furthermore, the lack of data on co-pay price precluded an analysis of the sensitivity of ADHD medication prescriptions to price.

Our small-area analysis cannot determine whether ADHD medication appears overused or underused in particular US counties. Indeed, we know of no clinical “gold standard” ADHD medication rate to which we can compare county-level results. Nevertheless, our findings highlight inequities in health care delivery across levels of physician supply and race/ethnicity. These inequities represent potential areas for improvement of the health care system. Systemwide policies that promote uniform standards of health care delivery may facilitate prudent allocation of health care resources.

Bibliography

- [1] Ablechild.org. State Legislation: Child Labeling & Drugging Bills & Resolutions Passed. Available at: <http://ablechild.org/slegislation.htm>. Accessed on July 10, 2010.
- [2] Acemoglu and Finkelstein, 2008, “Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector,” *Journal of Political Economy*, Vol. 115, No. 5, pp. 837 – 880.
- [3] [Anon.] Mental health in the United States: prevalence of diagnosis and medication treatment for Attention-Deficit/Hyperactivity Disorder—United States, 2003. *Morb Mortal Wkly Rep* 2005;54:842-847.
- [4] Bagwell CL, Molina BS, Pelham WE, Hoza B. Attention-deficit hyperactivity disorder and problems in peer relations: predictions from childhood to adolescence.” *J Am Acad Child Adolesc Psychiatry* 2001;40: 1285-92.
- [5] Barkley RA. 2003. Attention-Deficit Hyperactivity Disorder. In: Mash EJ and Barkley RA, eds. *Child Psychopathology: Second Edition*. New York: Guilford; 75-143.
- [6] Barkley RA, Fischer M, Edelbrock CS, et al. The adolescent outcome of hyperactive children diagnosed by research criteria: I. An 8-year prospective follow-up study. *J Am Acad Child Adolesc Psychiatry* 1990;29:546-57.
- [7] Benesch, Witter, Wilder, Duncan, Samsa, and Matchar, 1997. “Inaccuracy of the International CManaged Careclassification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease,” *Neurology*, Sept. 1, 1997, Vol. 49, No. 3, pp. 660 – 664.
- [8] Bhatara V, Feil M, Hoagwood K, et al. National trends in concomitant psychotropic medication with stimulants in pediatric visits: practice versus knowledge. *J Atten Disord* 2004;7:217-26.
- [9] Bokhari F, Mayes R, Scheffler RM. An analysis of the significant variation in psychostimulant use across the U.S. *Pharmacoepidemiol Drug Saf* 2005;14: 267- 75.

- [10] Brownell MD, Mayer T, Chateau D. The incidence of methylphenidate use by Canadian children: what is the impact of socioeconomic status and urban or rural residence? *Can J Psychiatry* 2006;51:847-54.
- [11] Bussing R, Schoenberg NE, Perwien AR. Knowledge and information about ADHD: evidence of cultural differences among African-American and white parents. *Soc Sci Med* 1998;46:919-28.
- [12] Bussing R, Schoenberg NE, Rogers KM, et al. Explanatory models of ADHD: Do they differ by ethnicity, child gender, or treatment status? *J Emot Behav Dis* 1998;6:233-242.ital accumulation: The case of ADHD. *J Health Economics* 2006; 25: 1094 - 1118.
- [13] Card, Dobkin, and Maestas, 2008. "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare," *American Economic Review*, Vol. 98, No. 5, pp. 2242 – 2258.
- [14] A CBO Paper, 2008. "Technological Change and the Growth of Health Care Spending," January 2008.
- [15] Chiang and Spat, 1982. "Imperfect Price Discrimination and Welfare," *Review of Economic Studies*, Vol. 49, No. 2, pp. 155 – 181.
- [16] Cohen P, Hesselbart CS. Demographic factors in the use of children's mental health services. *Am J Public Health* 1993;83: 49 - 52.
- [17] Coulam and Gaumer, 1991. "Medicare's Prospective Payment System: A Critical Appraisal," *Health Care Financing Review*, Annual Supplement, pp. 45 – 77.
- [18] Cox ER, Halloran DR, Homan SM. Trends in the prevalence of chronic medication use in children: 2002-2005. *Pediatrics* 2008;122: e1053-61.
- [19] Cox ER, Motheral BR, Henderson RR, et al. Geographic variation in the prevalence of stimulant medication use among children 5 to 14 years old: results from a commercially insured US sample. *Pediatrics* 2003;111: 237-43.
- [20] Currie J, Stabile M. Child mental health and human cap
- [21] Cutler, 1995. "The Incidence of Medical Outcomes Under Prospective Payment," *Econometrica*, Vol. 63, No. 1, pp. 29 – 50.
- [22] Eberstadt M. Why Ritalin rules. PoliCox ER, Motheral BR, Henderson RR, et al. Geographic variation in the prevalence of stimulant medication use among children 5 to 14 years old: results from a commercially insured US sample. *Pediatrics* 2003;111: 237-43. *cy Rev* 1999;94:24-44.

- [23] Education Week: Editorial Projects in Education, Research Center. Available at: <http://www.edweek.org/rc/2007/06/07/edcounts.html>. Accessed on March 3, 2008.
- [24] Ellis and McGuire, 1986. "Provider Behavior Under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, Vol. 5, pp. 129 – 152.
- [25] Feder, Hadley, and Zuckerman, 1987. "How Did Medicare's Prospective Payment System Affect Hospitals?" *New England Journal of Medicine*, Vol. 317, No. 14, pp. 867 – 873.
- [26] Fisher ES, Wennberg JE, Stukel TA, et al. Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. *Health Serv Res* 2000;34: 1351-62.
- [27] Ford T, Goodman R, Meltzer H The relative importance of child, family, school and neighbourhood correlates of childhood psychiatric disorder. *Soc Psychiatry Psychiatr Epidemiol* 2004;39:487-96.
- [28] Fulton BD, Scheffer RM, Hinshaw SP, et al. ADHD Diagnostic Prevalence and Medication Use Variation across U.S. States: An Examination of Health Care Providers and Education Policies. *Psychiatric Services* 2009;60:1075-83.
- [29] Gold, Chu, Felt, Harrington, and Lake, 1993. "Effects of Selected Cost-Containment Efforts: 1971 – 1993," *Health Care Financing Review*, Vol. 14, pp. 183 – 226.
- [30] Gottlobet et. al, 2001. "Medicare Hospital Prospective Payment System: How DRG Rates are Calculated and Updated," Office of Inspector General, Office of Evaluations and Inspections.
- [31] Harris, 1977. "The Internal Organization of Hospitals: Some Economic Implications," *Bell Journal of Economics*, Vol. 8, pp. 467 – 482.
- [32] Hinshaw SP. Is ADHD an impairing condition in childhood and adolescence? In: Jensen PS, Cooper JR, eds. *Attention-deficit hyperactivity disorder: State of the science, best practices*. Kingston, NJ: Civic Research Institute. 2002;5:15-21.
- [33] InterStudy Publications. *Trend Data Enrollment Count 2001 – 2008*. Nashville; Decision Resources, Inc. Updated November 11, 2008.
- [34] Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. New York: Wiley;1982:341-42.
- [35] MacIntyre, Ackland, Chandraraj, and Pilla, 2008, "Accuracy of ICD-9-CM codes in hospital morbidity data, Victoria: implications for public health research," *Australian and New Zealand Journal of Public Health*, Vol. 21, Issue 5, pp. 477 – 482.

- [36] Mayes R, Bagwell C, Erkulwater J. ADHD and the rise in stimulant use among children. *Harv Rev Psychiatry* 2008;16:151-66.
- [37] McCarthy et al. 2000. "Does Clinical Evidence Support ICD-9-CM Diagnosis Coding of Complications?" *Medical Care*, Vol. 38, Issue 8, pp. 868-876.
- [38] McClellan, 1996. "Medicare Reimbursement and Hospital Cost Growth," in David Wise (ed) *Advances in the Economics of Aging*. Chicago: Chicago University Press.
- [39] National Center for Educational Statistics Common Core of Data [database online]. Washington, D.C.; US Dept. of Education, Institute of Education Sciences. Updated March 10, 2008.
- [40] Parashar, Holman, Clarke, Bresee, and Glass. 1998. "Hospitalizations Associated with Rotavirus Diarrhea in the United States, 1993 through 1995: Surveillance Based on the New ICD-9-CM Rotavirus-Specific Diagnostic Code," *Journal of Infectious Disease*, Vol. 177, No. 1, pp. 13-17
- [41] Rappley MD. Clinical practice. Attention deficit-hyperactivity disorder. *N Engl J Med* 2005;352:165-73.
- [42] Rappley MD, Gardiner JC, Jetton JR, et al. The use of methylphenidate in Michigan. *Arch Pediatr Adolesc Med* 1995;149:675-9.
- [43] Ray GT, Levine P, Croen LA, Bokhari FA, Hu TW, Habel LA. Attentiondeficit/ hyperactivity disorder in children: excess costs before and after initial diagnosis and treatment cost differences by ethnicity. *Arch Pediatr Adolesc Med*. 2006;160:1063-9.
- [44] Reinganum, 1989. "The Timing of Innovation: Research, Development and Diffusion," *Handbook of Industrial Organization*, ed. Schmalensee and Willig.
- [45] Rowe R, Maughan B, Goodman R. Childhood psychiatric disorder and unintentional injury: findings from a national cohort study. *J Pediatr Psychol* 2004;29:119-30.
- [46] Scheffler RM, Hinshaw SP, Modrek S, et al. The global market for ADHD medications. *Health Aff (Millwood)* 2007;26:450-7.
- [47] Schneiders JM, Drukker J, van der Ende FC, et al. Neighbourhood socioeconomic disadvantage and behavioural problems from late childhood into early adolescence. *J Epidemiol Community Health* 2003;57:699-703.
- [48] Schneider H, Eisenberg D. Who receives a diagnosis of attention-deficit/ hyperactivity disorder in the United States elementary school population? *Pediatrics* 2006;117: e601-9.

- [49] Serden, Lindqvist, and Rosen, 2003. "Have DRG-based Prospective Payment Systems Influenced the Number of Secondary Diagnosis in Health Care Administrative Data?" *Health Policy*, Vol. 65, No. 2, pp. 101 – 107.
- [50] Singer, Natasha, 2009. "False Start on Laser Remedy for Fungus," *New York Times*, March 19.
- [51] Sloan, Morrissey and Valvona, 1988. "Medicare Prospective Payment and the Use of Medical Technologies in Hospitals," *Medical Care*, Vol. 26, No. 9, pp. 837 – 853.
- [52] Spanos B. Quotas, ARCOS, UN Report and Statistics. Conference Report: stimulant Use in the Treatment of ADHD. Washington, D.C.: Drug Enforcement Administration, Department of Justice; 1996.
- [53] Stevens J, Harman JS, Kelleher KJ. Race/ethnicity and insurance status as factors associated with ADHD treatment patterns. *J Child Adolesc Psychopharmacol* 2005;15:88-96.
- [54] Thomas CP, Conrad P, Casler R, et al. Trends in the use of psychotropic medications among adolescents, 1994 to 2001. *Psychiatr Serv* 2006;57:63-9.
- [55] US Census Bureau. People and Households Estimates [database online]. Available at: <http://www.census.gov/popest/estimates.html>. Accessed on March 8, 2008.
- [56] US Department of Health and Human Services, Health Resources and Services Administration [database online]. Health Care Professions 2001-2003. Available at: <http://www.arfsys.com/>. Accessed on March 3, 2008.
- [57] Wasserman RC, Kelleher KJ, Bocian A, et al. Identification of Bokhari F, Mayes R, Schefler RM. An analysis of the significant variation in psychostimulant use across the U.S. *Pharmacoepidemiol Drug Saf* 2005;14: 267- 75. of attentional and hyperactivity problems in primary care: a report from pediatric research in office settings and the ambulatory sentinel practice network. *Pediatrics* 1999;103: E38.
- [58] Weisbrod, 1991. "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care and Cost Containment," *Journal of Economic Literature*, Vol. 29, No. 2, pp. 523 – 552.
- [59] Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. *Bmj* 2002;325:961-4.
- [60] Wennberg J, Wennberg D. Practice variations and the use of prescription drugs. [Dartmouth Atlas of Health Care in Michigan web site]. Available at: <http://www.bcbsm.com/atlas/practice/shtml>. Accessed February 2, 2001

- [61] Williams JK, Klinepeter G, Palmes A, et al. Diagnosis and treatment of behavioral health disorders in pediatric practice. *Pediatrics* 2004;114:601-6
- [62] Zuvekas SH, Vitiello B, Norquist GS. Recent trends in stimulant medication use among U.S. children. *Am J Psychiatry* 2006;163:579-85.