Essays on Health Economics and Technology Adoption

by

Leila Agha

S. B. Massachusetts Institute of Technology (2006)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author .....
Department of Economics
May 15, 2011

Certified by .....
David Autor
Professor of Economics
Thesis Supervisor

Certified by .....
Michael Greenstone
3M Professor of Environmental Economics
Thesis Supervisor

Accepted by ...
Esther Duflo
Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics
Chairman, Department Committee on Graduate Theses

Essays on Health Economics and Technology Adoption

by

Leila Agha

Submitted to the Department of Economics
on May 15, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis studies the economics of technology adoption in the healthcare industry. The first chapter analyzes the impact of health information technology (HIT) on the quality and intensity of care delivered to Medicare inpatients. Building an organizational model, I show how the adoption of HIT may improve patient health and may either increase or decrease medical expenditures. Using Medicare claims data from 1998-2005, I estimate the effects of HIT by exploiting variation in hospitals' adoption statuses over time, analyzing 2.5 million inpatient admissions across 3900 hospitals. HIT is associated with an initial 1.3% increase in billed charges. Additionally, HIT adoption appears to have little impact on the quality of care, measured by patient mortality, medical complication rates, adverse drug events, and readmission rates. These results are robust to the addition of rich controls for pre-trends. The findings suggest that HIT is not associated with improvements in either the efficiency or quality of hospital care for Medicare patients, through five years after adoption.

In the second chapter, I investigate the scope for physician learning about the value and applications of new medical technologies across geographic regions. In particular, I analyze the diffusion of positron emission tomography and deep brain stimulation, using data on Medicare claims from 1998-2005. The mix of patient diagnoses treated with the new technologies changes substantially during the early stages of diffusion. Moreover, states that are late to adopt these technologies do not repeat the process of experimental learning undertaken by early adopters to discover which patients should receive the new treatment

In the third chapter, I analyze several policy initiatives that aim to manage the

usage of medical technologies and discuss key determinants of technology adoption that may be fruitful targets for future research and policy intervention. Effective technology policy must balance cost control with a recognition that new medical technologies have been associated with tremendous health and longevity gains. I find that existing Medicare coverage determinations and state certificate of need programs appear to have little influence on actual resource utilization, in part driven by lack of enforcement of existing policies.

Thesis Supervisor: David Autor
Title: Professor of Economics

Thesis Supervisor: Michael Greenstone
Title: 3M Professor of Environmental Economics

Acknowledgments

I am grateful to my advisors Michael Greenstone, David Autor, and Amy Finkelstein for their wisdom and guidance. Their teaching was invaluable as I developed my research interests and worked to complete this dissertation.

I also want to thank my classmates and other MIT faculty for their thoughtful comments throughout the past several years. I acknowledge many of them individually at the beginning of each chapter. In particular, Jason Abaluck, Tatyana Deryugina, Brigham Frandsen, Tal Gross, Danielle Li, Konrad Menzel, Amanda Pallais, and Christopher Palmer provided exceptional feedback.

Lastly, I would like to thank my husband Joshua Aronson, sisters Eva and Sindha Agha, and parents, Jennifer Cole and Gul Agha, for their love, patience and unwavering support. Joshua provided frequent, thoughtful consultations on the daily workings of the healthcare industry; his insight has enriched this work considerably.

Contents

Chapter 1

The Effects of Health Information Technology on the Costs and Quality of Medical Care [1]

1.1 Introduction

Technology adoption, and information technology in particular, has been linked to productivity growth in a wide variety of sectors. However, a historical perspective suggests caution is warranted in linking any particular technology to the promise of substantial, sustained productivity growth within a specific industry. Work by

[1]I would like to thank David Autor, Amy Finkelstein, and Michael Greenstone for invaluable guidance throughout this project. I gratefully acknowledge David Cutler and Frank Levy for their very helpful comments and suggestions. I thank Jason Abaluck, Joshua Angrist, Joshua Aronson, David Chan, Tatyana Deryugina, Joseph Doyle, Brigham Frandsen, Robert Gibbons, Jonathan Gruber, Nathan Hendren, Nancy Keating, Danielle Li, Amanda Pallais, Christopher Palmer, Michael Powell, Joseph Shapiro, Pian Shu, Catherine Tucker, Heidi Williams, and participants at MIT's labor lunch for their comments. I am also grateful to Mohan Ramanujan and Jean Roth for their assistance in obtaining and managing the data. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship.

the McKinsey Global Institute (2002) argues that the productivity acceleration of the 1990s, widely attributed to information technology (IT), was concentrated in a limited number of sectors, and IT was only one of several factors that combined to create the productivity jump.

In this paper, I analyze the impact of health information technology (HIT) on the costs and quality of medical care, testing whether the technology has demonstrated potential to improve the productivity of the health care sector. Health care costs were $2.3 trillion in 2008, and have been rising at a rate of 4.9% per year, in real terms, averaged over the past 40 years. As costs continue to grow, high rates of medical error and unnecessary or redundant expenditures persist. Many observers and policymakers are looking to health information technology as a key tool to improve the efficiency of the health care sector, by preventing medical errors, cutting redundant tests, and improving health outcomes. Estimates of potential savings due to HIT adoption vary, with widely cited, though controversial, work by the RAND Institute projecting a $142–$371 billion per year reduction in health spending (Hillestad et al. 2005).

Hospitals invested $4.7 billion in HIT in 2009, and are poised to increase their investments over the coming decade. The Health Information Management Systems Society estimates that hospitals will spend approximately $26 billion dollars in IT applications between 2010-2014 (HIMSS Analytics 2009). These expenditures will be driven partly by a federal program, the 2009 HITECH Act, which will implement reimbursement incentives and penalties designed to encourage HIT adoption. These new incentive payments are projected to increase net Medicare and Medicaid spending by $30 billion over nine years (2011-2019). However, the Congressional Budget Office estimates the total costs of the legislation to be markedly lower, $19 billion, since it predicts that HIT will reduce medical expenditures and thus reduce related

federal spending.

This study focuses primarily on two types of health information technology: electronic medical records and clinical decision support. Electronic medical records maintain patient information and physician notes in a computerized data base rather than a paper chart. Electronic records allow the provider to track the patient's health over time, read the input of other consulting physicians, or recall his own clinical assessment from a previous day or hospital visit. Clinical decision support provides timely reminders and suggestions to medical practitioners. Decision support may recommend screening tests based on a patient's age and medical conditions, flag drug-drug interactions and drug allergy information, or discourage the provider from repeating a test or imaging procedure by highlighting a previously ordered result.

The paper introduces a theoretical model of hospital organization that analyzes the potential effects of HIT adoption. Firstly, electronic records will raise the quality of communication across providers, which may in turn increase reliance on specialists and reduce redundant testing. Secondly, electronic records may reduce the effort cost of following a patient's treatment course, which may increase the intensity of provided treatment. Lastly, clinical decision support may encourage the primary physician to provide treatment for conditions that are less familiar to him, reducing reliance on specialists. The net impact of these three channels on total medical expenditures, health outcomes, and quality of care is ambiguous.

I use the theoretical model to inform an empirical analysis of the the impact of HIT, using Medicare claims data. HIT is associated with 1.3% higher medical expenditures, significant at the 10% level. The marked increase in the log of medical expenditures can be seen graphically in Figure 1, which plots coefficients from a nonparametric event study regression. HIT is adopted in year 0, and expenditure growth is seen after adoption. Other results find that length of stay and number

of physicians consulted do not change significantly after adoption. Despite the cost increases, HIT is associated with very modest reductions in patient mortality of 0.03 *percentage points* [95% confidence interval: -0.36 to 0.30 percentage points]. Further, there are no significant improvements in the complication rate, adverse drug events, readmission rate, or frequency of outpatient cancer screenings, after HIT adoption.

The results fail to measure a social benefit to HIT adoption over this period, although it should be noted that the finding is local both to the types of software systems commonly implemented over the study period, from 1998-2005, and the organizational structure of adopting hospitals. I will discuss these limitations further in the penultimate section of the paper.

These findings are estimated in a 20% sample of Medicare claims from 1998-2005; the sample includes 2.5 million inpatient admissions at 3880 hospitals. The claims data allows detailed tracking of patients' health outcomes, services rendered, and medical expenditures. HIT adoption is measured at the hospital level from the Health Information and Management Systems Survey (HIMSS).

A fixed effects econometric model exploits within-hospital across-time variation in HIT adoption status to estimate the effects of adoption. The multi-year panel data along with variation in the timing of HIT adoption, allows the inclusion of rich controls for time trends beyond those used in conventional difference-in-differences analysis; in particular, I control for state-year fixed effects, adopter-specific time trends, and differential trends that vary according to a hospital's baseline characteristics. I analyze potential threats to validity, testing for simultaneous changes in other hospital investments and probing the robustness of the results to any changes in patient sorting across hospitals.

This analysis has several advantages over previous research. First, it estimates the impact of HIT over a broad, national sample of hospitals, rather than presenting

a case study of a single institution or HMO (cf. Bates et al. 1999; Demakis et al. 2000; Evans et al. 1994; Javitt et al.). Second, it uses panel data to implement a difference-in-differences strategy, instead of relying on cross-sectional evidence (cf. DesRoches et al. 2010, Himmelstein et al. 2010). Third, in contrast to earlier work, the analysis brings together a large complement of outcome variables that include mortality outcomes, quality of care measures, and medical expenditures (cf. Miller and Tucker 2009, Cullough et al. 2010, Furukawa et al. 2010). Lastly, I implement a more robust empirical strategy that controls for a rich set of state-by-year fixed effects and differential time trends that vary by hospital characteristics, rather than imposing uniform time trends across hospitals (cf. Miller and Tucker 2009, Cullough et al. 2010).

The paper proceeds as follows. Section 2 provides a conceptual framework for how HIT adoption may affect the delivery of medical care. Section 3 describes the data in more detail, discusses the HIT adoption decision, and compares the characteristics of adopters and non-adopters. Section 4 presents the empirical strategy and results. Section 5 analyzes the policy implications and interpretation of these findings. The final section summarizes the results and concludes.

1.2 Conceptual Framework

Over the course of caring for a patient, a physician must make decisions about whether to deploy a variety of medical resources, including diagnostic tests, imaging, and specialized consulting physicians. On the one hand, each additional test or consultation has some chance of improving the patient's diagnosis, treatment, and ultimate health status. On the other hand, additional tests and consultations are costly, and poorly coordinated plans from specialists increase the risk of medical

complications and errors. When deciding which tests to order and specialists to call, a physician must carefully balance the quality of his existing information set, the additional costs, and the benefits and risks to the patient's health. HIT has the potential to change each of these tradeoffs, thus significantly shifting the practice of medicine, as well as its associated costs and expected health outcomes. The theory presented here lays out the basic mechanisms through which HIT may change the use of various inputs to healthcare production, and links these behavioral changes to testable empirical predictions. To date, there has not been a careful modeling of the impact of HIT on medical practices.

1.2.1 Modeling Physician Behavior and Patient Health

The existing public health literature on HIT predicts three main channels through which technology adoption may impact patient outcomes. First, electronic medical records are predicted to reduce costs by reducing redundant testing, and improve health outcomes by minimizing coordination failures. These two effects both work through the channel of improving communication across providers. Secondly, electronic medical records are thought to make it easier for the physician to track the patient's health needs and inputs over time, by reducing the effort cost of following a patient's care. Lastly, clinical decision support is designed to improve the diagnosis and treatment of patients by providing timely reminders and information about medical best practices. Each of these three channels—improved communication, lower effort costs of treating a patient, and better information about optimal care—could have other effects on the costs and quality of care provision beyond the conventional wisdom outlined above. The model I lay out in this section elucidates how changing these three parameters may affect the quality and costs of care.

I develop an organizational model of care delivery in an inpatient hospital setting, adapting the Dessein and Santos (2006) model to the particulars of the health production context. This work builds on the theoretical insight of Becker and Murphy (1992) that a key limitation to specialization is imposed by the cost of coordinating workers. It enriches the Becker and Murphy framework by modeling the specific tradeoffs within the healthcare production function associated with consulting a specialized physician, i.e. the potential for improved adaptation to the patient's idiosyncratic needs versus the increased risk of coordination failure across physicians.

A patient's health depends on how well his treatments match his medical needs. Suppose a patient is admitted to the hospital with two diagnoses that are contributing to an acute medical problem. The patient has idiosyncratic needs θ_1 for the primary diagnosis and θ_2 for the secondary diagnosis. For example, a patient with a hip fracture may be admitted to the orthopedic service for surgery; however, the patient may also suffer from heart disease and be medicated regularly with blood thinners to reduce the risk of heart attack. The surgeons will need to consider the patient's cardiac health when deciding whether he is fit for surgery, and the patient's usual blood thinners may need to be discontinued to minimize the risk of uncontrolled bleeding during surgery.

The patient's needs are indexed by the intensity of the optimal treatments, where higher values of θ correspond to more intensive use of medical resources. I assume that excessive treatment and insufficient treatment are both harmful to the patient's health. An insufficiently treated disease is likely to cause future health problems and prevent the patient from maintaining his optimal health. On the other hand, excessive treatment has its own set of health risks, e.g. higher risk of hospital acquired infection with each additional day spent as an inpatient, side effects and potential adverse reactions to drugs, and risks associated with surgical intervention.

The physician must offer a treatment plan for each diagnosis, which consists of a main treatment t^i; i indexes the problem being treated, $i \in \{1, 2\}$. The selected treatment will maximize the patient's health when it is exactly equal to the patient's need for that diagnosis, i.e. $t^i = \theta_i$.

In addition to selecting the appropriate treatment intensity for each of the patient's diagnoses, care providers must ensure that the treatment plans are harmonized across diagnoses. The patient's health may be compromised if the treatment of the primary condition is not coordinated with the treatment of the secondary condition. A third action, t^{12} is required to coordinate the treatments; the risk of negative treatment synergies is minimized when $t^{12} = t^2$. For example, preparation for surgery may require changing the management of a patient's chronic illnesses, e.g. pausing the prescription of blood thinners to prevent excessive bleeding during an operation; or, drugs prescribed by two different providers may have negative synergies which require adjustment of the patient's prescriptions.

Patient's health can be expressed mathematically by the following healthcare production function:

$$H = H^* - \phi[(t^1 - \theta^1)^2 + (t^2 - \theta^2)^2] - \beta[(t^{12} - t^2)^2] \tag{1.1}$$

In equation (1), H^* represents the patient's optimal health, ϕ parameterizes the importance of adapting the main treatment intensity to the patient's needs, and β parameterizes the importance of coordinating the treatment of the patient's two medical conditions.

Medical care is associated with both financial and effort costs. More intensive treatments require more effort from the physician, who will need to follow the patient more carefully, and are associated with higher medical expenditures. The physician

may choose to consider these increased financial and effort costs when selecting a treatment. I assume that both effort and financial costs are proportional to treatment intensity, t^i.

The physician will select a treatment plan to maximize his utility, which depends on the patient's health and the associated costs of treatment.

$$\begin{aligned} U &= \underbrace{\mathbb{E}[H^* - \phi[(t^1 - \theta^1)^2 + (t^2 - \theta^2)^2] - \beta(t^{12} - t^2)^2]}_{\text{expectation of patient health}} \\ &\quad -\gamma \underbrace{[\lambda t^1 + \lambda t^2}_{\text{treatment costs}} + \underbrace{1\{d = s\}(\alpha + \delta(1 - c))]}_{\text{specialist costs}} \end{aligned} \tag{1.2}$$

The parameter γ measures the relative weight put on costs in the physician's utility function, and λ is equal to the sum of effort costs plus financial costs associated with one unit of treatment. $1\{d = s\}$ is a indicator variable which equals one if a specialist is called. In the event that a specialist is called, monetary costs of care will increase by a fixed amount to reimburse the specialist and the primary physician must endure a fixed effort cost associated with requesting the specialist consultation. The sum of these two costs is paramaterized by α. In addition, the specialist will have to repeat δ worth of testing and imaging if he does not receive the information from the primary physician, which occurs with probability $1 - c$.

Financial incentives can either discourage or encourage intensive treatment in this model. Under capitation payments for Medicare inpatients, hospitals will generally prefer to reduce treatment intensity to reduce costs, i.e. financial incentives will make λ positive. On the other hand, for outpatient care and professional services not included in the capitation payments, financial incentives may encourage the physician to provide more intensive treatments, in order to increase his reimbursements. In this case, λ may be negative, and effort costs and financial incentives will have opposing

effects.

A key challenge to delivering optimal care is determining the best treatment course for a given patient. The population distribution of medical needs is $\theta_i \sim \mathrm{N}(\hat{\theta}, \sigma_\theta^2)$. A doctor cannot perfectly observe an individual patient's needs, and must rely on a noisy signal, $\tilde{\theta^i} = \theta^i + \epsilon_i$. The noise in this signal, ϵ_i is Normally distributed with a mean of zero, and a variance that depends on whether the treating physician decides to call a specialist.

A specialist is better at adapting to the patient's needs for the second diagnosis, because he observes a more precise signal of θ_2, i.e. $\epsilon_2 \sim \mathrm{N}(0, \sigma^2_{\epsilon,d})$, and $\sigma^2_{\epsilon,p} > \sigma^2_{\epsilon,s}$, where $d = p$ if the primary physician treats the patient, and $d = s$ if the specialized doctor treats the patient. This assumption is consistent with recent evidence that specialists obtain better health outcomes than general practitioners within their area of speciality (Ayanian et al. 2002, Landon et al. 2003, Wells and Sturm 1995).

Consulting a specialist will also introduce a risk of coordination failure, in the event that the two physicians do not successfully communicate their treatment plans to each other. As the reliance on specialized physician services has grown over the recent decades, there has been concern that coordination failures are increasing the frequency of medical errors (Ghandi et al. 2000, Epstein 2005). In the model, I parameterize the quality of communication with c; with probability $1-c$, the primary physician will not learn the specialist's treatment plan t^2. As a result, the primary physician will not be able to select the optimal coordinating action, i.e. $t^{12} \neq t^2$, which will reduce the patient's health.

There is significant regional variation in the reliance on medical specialists. Baiker and Chandra (2004) document that a one standard deviation increase in the supply of medical specialists is associated with 50% more Medicare beneficiaries seeing more than 10 doctors in the last six months of life. This evidence suggests that the decision

to call a specialist is a margin along which the primary physician may adjust his behavior, and there may indeed be scope for increasing or decreasing the reliance on specialists.

If the primary physician chooses not to call a specialist, he will choose t^1 and t^2 to maximize his utility as written in Equation (2), using the noisy signal he received of θ_1 and θ_2. Solving the maximization problem yields the optimal treatment:

$$t_p^{ii} = \frac{\frac{1}{\sigma_{\epsilon,p}^2}\tilde{\theta}^i + \frac{1}{\sigma_\theta^2}\hat{\theta}^i}{\frac{1}{\sigma_{\epsilon,p}^2} + \frac{1}{\sigma_\theta^2}} - \frac{\gamma}{2\phi}\lambda, \; i \in \{1,2\} \tag{1.3}$$

The first term in the treatment expression above is the patient's expected optimal treatment, given the signal of his medical needs and the population distribution of needs. As long as λ is positive, e.g. when the hospital receives capitation payments, this equation indicates that the offered treatment is less intensive than the unconstrained optimum due to the costs of care.

If no specialist is called, then there is no risk of communication failure, and the treatments for the patients' two conditions are perfectly coordinated, i.e. $t_p^{12} = t_p^{22}$. With the treatment in Equation (3) above, the patient's expected health status is:

$$H_p = H^* - 2\phi\frac{\sigma_\theta^2\sigma_{\epsilon,p}^2}{\sigma_\theta^2 + \sigma_{\epsilon,p}^2} - \frac{\lambda\gamma}{2\phi} \tag{1.4}$$

The noisier the signal of the patient's health, $\sigma_{\epsilon,p}^2$, and the higher the variance in the population distribution, σ_θ^2, the more difficult it will be for the physician to select the optimal treatment. In addition, higher costs of care are associated with lower overall health, since the patient will receive a lower than optimal treatment intensity.

I assume that the physician decides whether to call a specialist before evaluating the patient. If a specialist is called, the primary physician will not analyze the

patient's second medical problem, and so will not receive any signal of θ_2. If a specialist is called, the optimal treatment for medical problem 1 is the same as above, but the specialist's choice of treatment for medical problem 2 is now:

$$t_s^{22} = \frac{\phi}{\phi + \beta(1-c)} \frac{\frac{1}{\sigma^2_{\epsilon,s}}\tilde{\theta}^2 + \frac{1}{\sigma^2_{\theta}}\hat{\theta}^2}{\frac{1}{\sigma^2_{\epsilon,s}} + \frac{1}{\sigma^2_{\theta}}} + \frac{\beta(1-c)}{\phi + \beta(1-c)}\hat{\theta}^2 - \frac{\lambda\gamma}{2\phi} \tag{1.5}$$

The treatment is again related to the weighted average of the patient's signal $\tilde{\theta}^2$, and the average value $\hat{\theta}^2$; however, the weights are different from those in the single-physician case. On the one hand, the specialist observes a more precise signal of the patient's needs, which will tend to increase the weight he puts on the signal $\tilde{\theta}^2$. On the other hand, in the event of a communication failure, the primary physician will not learn about the specialist's treatment plan and will set the coordinating action $t_p^{12} = \hat{\theta}^2$, inducing coordination costs. The risk of this communication failure causes the specialist to tilt his action closer to the population mean $\hat{\theta}^2$ than he would in the case of perfect communication.

If a specialist is called, the patient's expected health status is:

$$H_s = H^* - \phi\frac{\sigma^2_{\theta}\sigma^2_{\epsilon,p}}{\sigma^2_{\theta} + \sigma^2_{\epsilon,p}} - \phi\left[\frac{\beta(1-c)}{\phi + \beta(1-c)}\sigma^2_{\theta} + \frac{\phi}{\phi + \beta(1-c)}\frac{\sigma^2_{\theta}\sigma^2_{\epsilon,s}}{\sigma^2_{\theta} + \sigma^2_{\epsilon,s}}\right] - \frac{\lambda\gamma}{2\phi}. \tag{1.6}$$

Now, the patient's health status depends not only on the precision of his health signal and the cost of his care, but also on the quality of communication between his two providers.

The primary physician will choose to call a specialist if doing so will improve the patient's expected health by a large enough margin to compensate for the increased costs. Using the expressions for the optimal treatments written above, and comparing the patient's health production function under the specialist and no-specialist

scenarios, I derive that the primary physician will call a specialist if and only if the following condition is met:

$$\frac{(\sigma^2_{\epsilon,p} - \sigma^2_{\epsilon,s})}{\sigma^2_\theta} \frac{(\sigma^2_\theta + \sigma^2_{\epsilon,s})}{(\sigma^2_\theta + \sigma^2_{\epsilon,p})} > \frac{\beta(1-c)}{\phi + \beta(1-c)} + \gamma 1\{d = s\}(\alpha + \delta(1-c)) \tag{1.7}$$

The primary physician will call a specialist if the gains from improved adaptation outweigh the additional billing costs and the risk of coordination costs in the event of communication failure. As the gap between the precision of the primary physician's signal and the specialist's signal grows, the primary physician becomes more likely to call a specialist. As the communication channel worsens, the physician becomes decreasingly likely to request a consultation.

1.2.2 Modeling the Effects of HIT

HIT has the potential to change three key parameters of the above model, ultimately impacting both the costs and the quality of delivered care.

Proposition 1. *Electronic patient records raise c, the quality of communication. Improved communication may increase reliance on specialists, while reducing redundant testing. It will improve health outcomes and may either increase or decrease medical expenditures.*

Electronic medical records allow the primary physician to access the specialist's report easily and in a timely manner. Thus, electronic records may improve the quality of communication between the specialist and the primary physician, raising the parameter c .

In Equation (7), increases in c will increase the likelihood that a specialist is called by reducing the risk of coordination failure and reducing costs associated with

redundant testing. Calling a specialist will reduce adaptation costs, since the specialist is better at adapting to the patient's needs. As a result of this increased reliance on specialists, patient health should improve. In the empirical work, I measure patient health by the mortality rate, which should weakly decline after the adoption of electronic records. Medical complications, adverse drug events, and readmission rates may also decrease, since these may be signals of poorly coordinated care plans.

Improving the communication channel will reduce the risk of redundant testing, conditional on a specialist being called. Existing evidence suggests that redundant testing is rampant in both the inpatient and outpatient setting. One study estimated 2.7% of inpatient hospital spending is on redundant tests (Jha et al. 2009), and another found that in the three months following an admission, 8.6% of follow-up tests were redundant (Bates et al. 1998). By making it easier to access images and notes produced by other physicians, HIT may reduce the frequency of redundant testing. In the empirical work, I will directly test whether the frequency of repeated imaging falls after the introduction of an electronic medical record.

The net impact of the improved communication channel on total medical expenditures is ambiguous. If improved communication induces more reliance on specialists, this will tend to increase billing since the specialists will submit additional charges for their evaluation and management of the patient. In cases where a specialist was being consulted before the introduction of electronic records, redundant testing should decrease as a result of the improvement in communication quality. The drop in redundant testing will tend to reduce billing. The empirical results will test which effect dominates.

Proposition 2. *Electronic records may reduce the effort cost of intensive treatment,* λ. *This will increase the treatment intensity. Under a capitation payment system,*

an increase in treatment intensity is likely to improve patient health.

A second effect of electronic records may be to reduce the effort cost of treating a patient more intensively. Monitoring the patient and following up on additional testing and new treatment plans may be easier if relevant information is consolidated into a single electronic system. For example, the physician may be more likely to order one more set of imaging to rule out a possible but unlikely diagnosis, if the process of ordering and viewing the result is made easier.

A decrease in λ will increase the treatment intensity for both diagnoses, regardless of whether a specialist is consulted. For this reason, overall medical expenditures may rise after HIT adoption, driven by longer lengths of stay, or higher expenditures on diagnostic testing, medications, and surgical treatment. I will test for expenditure increases in each of these categories. When financial incentives associated with capitation payments drive the treatment intensity to be lower than then health-maximizing optimum, then this reduction in λ will improve patient health by increasing treatment intensity. On the other hand, in a fee-for-service setting, there may be over-provision of medical treatment before HIT adoption; HIT could further amplify this mechanism, thus reducing patient health outcomes.

Proposition 3. *Clinical decision support may reduce the noise in the primary physician's signal of the patient's needs $\sigma^2_{\epsilon,p}$. This will reduce the likelihood that a specialist is called and reduce medical expenditures. The total impact on patient health is technically ambiguous, but it is likely that improvements in health would dominate.*

Clinical decision support software may have different effects on care delivery than electronic records. In particular, clinical decision support could significantly decrease the specialist's advantage, by making information about diagnosis and treatment of a

wide variety of conditions easily accessible to all care providers. The primary physician may feel reassured that the reminders and information available in the decision support software improve his ability to treat a patient with a medical condition which is less common or familiar to him. In the model, this is represented by a reduction in $\sigma^2_{\epsilon,p}$.

By improving the primary physicians ability to recognize the patient's optimal treatment, decision support will reduce the likelihood that a specialist is called. With fewer specialists consulted, decision support may reduce costly coordination failures, medical expenditures, and redundant testing. Patients' health outcomes could either improve or worsen after clinical decision support adoption. The physician may stop calling a specialist after decision support implementation to forgo the added costs of a specialist's workup, even if the patient's health would still be improved by the consultation. However, conditional on having the primary physician treat the condition, the patient's health should improve after the implementation of clinical decision support.

The propositions above highlight the effects of HIT through the three channels cited in the existing literature. I find that the impact of these changes on patient health and medical expenditures is theoretically ambiguous, contrary to conventional wisdom. Thus empirical work is critical to evaluating which impulse dominates in practice. I empirically test the net effect of these three channels by studying how the intensity of treatment, patient health, and quality of care evolve after HIT adoption.

1.3 Data and Descriptive Statistics

1.3.1 Data sources and sample construction

I study the impact of HIT on the costs and quality of care between 1998-2005, using data from three sources: Medicare Claims Data from the Center for Medicare and Medicaid Studies, the Health Information and Management Systems Survey (HIMSS) conducted by the Dorenfest Institute, and the American Hospital Association Annual Survey.

The HIMSS tracks HIT adoption at hospitals across the country; it includes questions about a wide variety of HIT functionalities and the timing of technology adoption. The annual survey includes 90% of non-profit, 90% of for-profit, and 50% of government-owned (non-federal) hospitals. The survey excludes hospitals with fewer than 100 beds. I construct an indicator variable of HIT adoption which equals one if the hospital has contracted either clinical decision support or computerized patient records. As reported in Panel A of Table 1, 54% of hospitals have contracted at least one of these two technologies by 1998, and an additional 23% of hospitals contract HIT for the first time during the study period.

THe HIMSS data is, to my knowledge, the only broad panel data on HIT adoption over this period. A shortcoming of the data is that although it differentiates the adoption of many different electronic capabilities, it does not record information on the quality of the HIT systems or the precise functionalities they include. I turn to the 2008 survey conducted by the American Hospital Association, reported by Jha et al. (2009a; 2009b), to understand which specific capabilities are likely to be included in the HIT installations I observe. This smaller survey covers 2370 hospitals, as compared to the 3880 hospitals included in the broader HIMSS data, and provides

a snapshot of HIT installations in the 2008 survey year, a few years after the end of my study period in 2005.

Jha et al. (2009b) report that the four most common components of electronic patient records are demographic characteristics (fully implemented in one or more unit at 89% of surveyed hospitals), medication lists (68%), discharge summaries (66%), and list of current medical conditions (48.5%); these four functionalities are highly likely to be features of the electronic record systems I observe. Daily physician progress notes (separate from discharge summaries) are less broadly diffused (29%), but may be a key component of about half of the record systems. Thus, electronic record systems allow physicians to easily track a patient's course of medications, treatment course during previous hospital visits, and current medical needs.

The most common feature of clinical decision support are drug allergy alerts (fully implemented in at least one unit at 68% of surveyed hospitals) and drug-drug interaction alerts (68%). Roughly half of the clinical decision support systems includes clinical guidelines and reminders, such as reminders to prescribe beta blockers after a myocardial infarction (implemented at 30% of surveyed hospitals) or provide pneumonia vaccines (38%). Decision support should improve the treatments primarily by ensuring that prescriptions are appropriately targeted and consistent with current standard of care guidelines.

I link the HIT adoption survey to data on all Part A and Part B Medicare claims for a 20% sample of patients over eight years, from 1998-2005. The Medicare claims data allows me to construct measures of patient health, medical expenditures, and the quality of hospital care. Because HIT adoption is observed at the hospital-level, I cannot observe which outpatient care providers are linked to an interoperable HIT system. For this reason, my analysis focuses on patients receiving inpatient care. The sample includes patients admitted to a hospital with a primary diagnosis of acute

myocardial infarction, stroke, hip fracture, lung cancer, colon cancer, gastrointestinal hemorrhage, or pneumonia. This set of diseases was chosen following previous work such as Baiker and Chandra (2004) because hospitalization for these conditions is likely to be a good proxy for disease incidence. I follow all inpatient and outpatient Medicare claims for these patients for one year following their first in-sample hospital admission.

In addition to studying this inpatient population, I track the use of outpatient preventive care services performed in the full 20% sample of Medicare recipients. For the study of outpatient preventive care, I analyze the subset of services performed by institutional providers, i.e. hospitals, for which I observe the HIT adoption status of the institution. I focus on mammography and colon cancer screenings, as these services are commonly provided on an outpatient basis within a hospital setting.

Limiting the primary analysis to Medicare patients does require compromising the breadth of the population studied. If HIT has heterogeneous effects which depend on the patient age group, then a limitation of this analysis is that it only identifies effects on the elderly population. The benefit of using Medicare data is the rich set of claims observable in this data allows me to measure the effects of HIT on a broad range of relevant outcome variables, in a panel data setting. In addition, Medicare enrolled 15% of the US population and accounted for 20% of total health spending in 2007, fractions that are likely to grow as the population ages. Lastly, elderly patients are highly likely to have multiple medical problems, putting them at greater risk for the coordination failures and mistakes that health IT is specifically designed to prevent.

Lastly, I complement the Medicare claims and HIMSS data with data from the annual American Hospital Association (AHA) survey. The AHA survey allows me to measure several key hospital characteristics, including hospital investments in other

diagnostic and therapeutic technologies, staffing levels, and total number of patient admissions.

Data are matched across these three sources using the hospital's Medicare provider number. Patients are indicated as exposed to HIT according to the adoption status of the admitting hospital, where they received inpatient treatment. HIT adoption status is observed for a sample of 3880 hospitals. I was able to match 90% of Medicare inpatient stays to the IT adoption status of the admitting hospital. There are a total of 2.5 million individuals in the inpatient sample, each of whom is tracked for one year following their hospital admission.

1.3.2 Summary Statistics

Table 1 provides an overview of hospitals' 1998 baseline characteristics by their adoption status. Adopting hospitals are larger on average than non-adopters, with twice as many inpatient beds, and an average of 8300 annual admissions compared to 3300 admissions for non-adopters. Adopters are more likely to be academic hospitals, be designated as a trauma center, and have adopted PET, MRI, and CT scanners. New scanners and new HIT systems both require large fixed cost investments, which may be more profitable for larger hospitals.

Patient characteristics do not differ as dramatically across hospitals. Comparing columns (1) and (2), adopters serve a slightly younger and more racially diverse population. Consistent with the younger population, adopters have a 0.7 percentage point lower one-year mortality rate amongst in-sample patients, as reported in Panel D. Total medical expenditures in the one year following an inpatient admission are 30% higher for patients at adopting hospitals.

Adopting hospitals offer more intensive treatment in the pre-period along a num-

ber or margins, including more ordering of electrocardiograms, longer hospital stays, and more physicians evaluating each patient. In addition, adopters have notably higher self-reported complication rates, with at least one code indicating a complication reported for 6.5% of patients, as compared to 4.4% of patients amongst non-adopting hospitals. By contrast, the rate of self-reported adverse drug events is not higher amongst adopting hospitals.

The pre-period differences between adopters and non-adopters suggest that it will be important in the remaining analysis to control for baseline differences and allow for the possibility of differential trends across hospitals with different adoption statuses. In addition, I will show that my results are robust to omitting the set of non-adopting hospitals from the estimation sample. The estimates become less precise when non-adopters are removed, so for this reason, I include them in the main results. I will further discuss the strategies I use to account for this heterogeneity in Section 4.1.

Two striking characteristics of this population indicate that it is particularly well-suited to identifying the impact of HIT. First, these patients are quite ill, with a 10% mortality rate in the baseline year. Thus, improvements in health may well be expected to occur along the margin of one year mortality, making survival a reasonable indicator of health in this sample. Second, the average patient in this sample sees over ten unique physicians during their admission and the year following. The large number of providers per patient suggests that there is significant scope for coordination failures within this population, if information is overlooked or missed as the patient shuttles between different doctors.

1.4 Empirical Results

1.4.1 Empirical Estimation Strategy

To examine how HIT affects medical expenditures and patient health, I use a fixed effects regression model as follows:

$$Y_{ht} = \alpha_h + \beta_1 HIT_{ht} + \gamma_{st} + X_{ht}\delta + \mu HITAdopter_h Yr_t + Q_h Yr_t \nu + \epsilon_{ht} \quad (1.8)$$

Y_{ht} is the outcome variable for a hospital h at time t. α_h are hospital fixed effects. HIT_{ht} is a binary variable equal to one if a hospital has contracted either a clinical decision support or an electronic medical records system in the current year or in an earlier year. γ_{st} is a vector of state-year fixed effects. X_{ht} is a vector of hospital and patient characteristics. I control for the hospital's investment in CT, MRI, and PET scans, as well as its status as a trauma hospital. Included patient characteristics are 1-year age bins, race, sex, and primary diagnosis. $HITAdopter_h$ is a dummy variable which equals one if the hospital has adopted HIT by the end of the study period in 2004; this variable is interacted with a linear time trend. Lastly, Q_h is a vector of hospital size dummy variables, indicating which quartile the hospital falls into according to number of inpatient admissions in the 1998 base year; these variables are also interacted with the time trend.

This specification is analogous to a difference-in-differences framework. The key coefficient of interest is β_1, which indicates how the outcome variable changes after a hospital has adopted health information technology. I compare the outcome variable within an adopting hospital before and after HIT adoption, controlling for the estimated counterfactual time trend the hospital would have experienced, had it not adopted HIT. Included state-year fixed effects capture state-specific shocks

and trends in medical practice patterns or unobserved characteristics of the patient population. Allowing unrestricted, differential trends by quartile of hospital size and the hospital's eventual adoption status allows for different types of hospitals to experience different trends.

Identification of equation 1.8 is based on the assumption that adoption of HIT is not coincident with other discontinuous changes in hospital organization, provider quality, or unobserved patient characteristics that would affect the measured outcome variables. Hospitals of the same size quartile, same eventual adoption status, and in the same state must be on parallel trends in the absence of HIT adoption, after controlling for observable changes in patient diagnoses and demographics.

In addition to including state*year fixed effects, I have also run every specification including county*year fixed effects, and the results do not change substantially. An F-test rejects the joint significance of the county-year fixed effects, after the inclusion of state-year effects, so county-year fixed effects are omitted from the specifications reported here.

Observations are at the hospital-year level based on the annual average of each variable across all in-sample patients admitted to that hospital. Accordingly, observations are weighted by the number of in-sample patients. There are 27,317 observations in total. Standard errors are clustered at the hospital level.

I report a second set of results which estimates a full trend-break model, as follows:

$$Y_{ht} = \alpha_h + \beta_1 HIT_{ht} + \beta_2 HIT_{ht} Yr_t + \gamma_{st} + X_{ht}\delta + \mu HITAdopter_h Yr_t + Q_h Yr_t \nu + \epsilon_{ht} \tag{1.9}$$

The regression above includes both β_1 and β_2 terms, allowing not only for a shift in the outcome variable at the time of HIT adoption, but also a break in the time trend.

If HIT adoption gradually changes hospital behaviors each year after adoption, then estimating β_2, the trend break, will allow more accurate estimates of the full effects of HIT.

In addition to reporting the two sets of regression estimates described above, a related set of graphical results are included. These graphs are based on regressions which include the same set of fixed effects and controls listed in Equation 1.8, but replace the key independent variable with a series of dummy variables indicating the year in normalized time. The coefficients on these normalized year dummy variables provide a year-by-year estimate of the treatment effect in event time.

This series of graphs provides a nonparametric way of visually assessing how the outcome variable evolves after HIT adoption. The HIMSS survey data measures the year in which HIT was first contracted from the software vendor; installation and implementation may be rolled out gradually in the year or two following the initial contract. Thus, these figures are useful for assessing whether the full impact of HIT is not realized until a few years after adoption.

Each of these specifications remains vulnerable to the possibility that some unobserved characteristic of the hospital or its patients changed right at the time of HIT adoption, thus confounding the estimated treatment effects. I deal with this threat to validity in three ways. First, I directly control for observed patient and hospital characteristics that may be evolving at the time of HIT adoption. Second, in Section 1.4.5, I demonstrate that these observable characteristics are not changing discontinuously at the time of HIT adoption. However, both of these approaches are vulnerable to the possibility that *unobserved* patient characteristics are changing at the time of HIT adoption. To address selection on unobservables, I present results that aggregate the unit of observation from the hospital to the county to account for the possibility that patient sorting may be more severe across hospitals within a

county, rather than across counties. I show that the conclusions do not change in the county-aggregated specifications.

Lastly, to improve the power of my tests and reduce the rate of false positive results, I group outcome variables into three conceptual categories and create standardized effect measures across these outcomes. The three domains are: intensity of treatment, quality of inpatient hospital care, and the frequency of outpatient preventive screening. These groupings allow me to perform omnibus tests analyzing whether HIT is affecting treatment patterns in a particular direction within a domain. I report both separate results for each outcome variable, as well as the aggregated standardized effect. I account for the cross-equation covariance structure of the error terms when estimating standard errors for each outcome within a domain. Standard errors remain clustered at the hospital level.

The standardized effect is constructed by combining the estimated coefficients across each outcome variable within a domain. In particular, the standardized effect equals:

$$\sum_{j \in J} \frac{1}{J} \frac{\beta_{ij}}{\sigma_j}, , i \in \{1, 2\} \tag{1.10}$$

where β_{1j} is estimated by Equation 1.8 for outcome variable j, or similarly for both β_{1j} and β_{2j} estimated by Equation 1.9. σ_j is the standard deviation of the outcome j amongst the hospitals that eventually adopt HIT, in the baseline year of 1998, prior to their adoption. Dividing by the standard deviation harmonizes the units across the diverse outcome variables. J is the total number of outcomes within a domain.

1.4.2 Impact of HIT on Mortality and Expenditures

Table 2 reports the results on patient health and medical expenditures. Columns (1) and (3) report results from estimating Equation 1.8, columns (2) and (4) report

the full trend-break model specified in Equation 1.9. Because most outcomes do not display significant breaks in slope of the time trend after HIT adoption, I focus on the Equation 1.8 specification throughout the discussion.

The relationship between medical expenditures and HIT is presented in columns (1) and (2) of Table 2, as well as in Figure 1. HIT is associated with initial increases in spending of around 1.3% (p=5.6%), in column (1). The 95% confidence interval suggests that there are no substantial decreases in expenditure, with the lower-bound at a 0.03% decrease and the upper bound a 2.6% increase. Column (2) estimates a larger, significant initial increase in expenditure of 1.8%, but suggests that adopters eventually return to the baseline trend, although the change in the trend is not statistically significant. The three-year effect is reported at the bottom of the table for the trend-break specifications. The three-year effect is within my observation sample for most adopting hospitals, so I focus on that estimate when analyzing the trend-break model. In column (2), I find that three years after HIT adoption, the increase in expenditures is estimated to be 0.8%, although a t-test fails to reject the null hypothesis of no change in expenditures.

Figure 1 illustrates the expenditure result graphically. The trend in expenditures appears to accelerate in the few years immediately following adoption, although the acceleration in the trend is largely within the bounds of the estimated 95% confidence interval. I will further unpack the relationship between HIT adoption and medical expenditures in Section 1.4.3, analyzing which services drive the estimated increase in expenditure.

Results reported in Table 2, columns (3) and (4), find no significant relationship between HIT adoption and 1-year patient mortality, contrary to the predictions of many observers. The point estimate in Table 2, column (3) suggests that HIT is associated with a 0.03 percentage point reduction in the mortality rate. The 95%

confidence interval on the mortality effect in Table 2, column (3), bounds an effect not larger in magnitude than a decrease of 4 deaths or increase of 3 deaths per 1000 patients, relative to a mean of 100 deaths per 1000. Recall that this is an acutely ill, elderly population, so the baseline mortality rate is quite high. Figure 2 confirms the small, insignificant effect size, with the mortality rate in years 0 through 5 remaining very close to the baseline levels before HIT adoption.

The modest increases in medical expenditures, coupled with the lack of significant improvement in the mortality rate, suggests further evidence of flat-of-the-curve medicine. By reducing the effort cost of intensive treatment, HIT may encourage the provision of care, even if the medical returns to this additional care are low. The lack of a mortality response to the expenditure increase is consistent with evidence from many recent studies (Murphy and Topel 2003; Baiker and Chandra 2004). An alternative explanation of the increased expenditures is that HIT systems improve billing capture, without changing medical behavior. This behavior would also drive increased expenditures with no commensurate improvement in mortality rates.

In sum, although I find no evidence of cost savings or substantial mortality benefit, the cost and mortality estimates cannot exclude the possibility that HIT is a valuable investment. To further investigate the potential impact of HIT, I analyze a number of different measures of how HIT may change the intensity of treatment and quality of care provided. If HIT is to improve health, it may do so through a number of channels previously outlined: reducing adverse drug events, medical complications, and readmissions. If HIT is associated with cost savings, we may observe shorter lengths of stay and less repeated imaging. Testing these specific channels will illuminate the cost/benefit tradeoffs by providing additional evidence of the impact along both of these margins.

1.4.3 Impact of HIT on the Intensity & Efficiency of Care

Medical Expenditures

The increase in medical expenditures reported in Table 2 could be driven by several factors: more screening and diagnostic tests, longer hospital visits, increased spending on treatment interventions such as medications and operations, or higher spending on outpatient care and physician services. Table 3 analyzes separately the relationship between HIT adoption and each of these categories of spending. I find that HIT is associated with temporary increases in three of the four categories of spending, but the magnitude and statistical significance of the increase varies considerably by category.

Spending on diagnostic testing and imaging increased by 1.6% following HIT adoption, or about $160 per patient, significant at the 10% level, as estimated in the column (1) specification. Figure 3 illustrates the trend expenditure growth on diagnostic testing, after HIT adoption. The trend break model estimated in Table 3, column (2), suggests that expenditure growth is slower amongst adopters after HIT adoption, although the coefficient is small and imprecisely estimated. This category of expenditure has the highest estimated 3-year effect, a 1.3% increase, although the 3-year estimate is not statistically distinguishable from zero. The estimated increase in diagnostic testing accords with the model prediction that HIT may reduce the effort costs of ordering and following additional tests, and, as a result, may increase a physician's propensity to order a more intensive work-up. The estimate suggests that any potential, unobserved decrease in redundant testing is counteracted by a larger overall increase in testing, and HIT is not associated with substantial cost savings due to less frequent imaging.

I also find evidence of increased expenditures on hospital stays, although the asso-

ciation is not statistically significant at conventional levels in either specifications. In particular, it appears that there is an initial bump in expenditures on inpatient care, followed by a gradual return to the baseline trend. The initial expenditure increase is about 1.1%, which is smaller than the estimated effect on diagnostic imaging. By three years post-adoption, inpatient expenditures are estimated to be only 0.7% higher amongst adopters, which is not statistically significant.

A similar pattern emerges for expenditures on medications, durable medical equipment, operating room staffing, and blood transfusions. This expenditure index, with results reported in Table 3, columns (5) and (6), measures spending on treatments (as opposed to diagnosis or monitoring). In the column (5) specification, I find an increase of 1.5% higher spending. Again, the full trend-break model indicates a gradual return to the baseline trend. After three years, expenditures on treatment interventions and medications are only 0.8% higher amongst HIT adopters, which is not statistically distinguishable from zero.

Physician services and outpatient care experience the smallest change in spending after HIT adoption. The point estimate in column (3) suggests a decrease in expenditure of 0.9%, but the effect is not statistically distinguishable from zero. This category of spending on professional services and outpatient care does not experience any of the expenditure growth found in other categories, and indeed, may slightly offset the other increases in expenditures.

In sum, there is evidence of increased spending associated with more intensive diagnostic work-ups and interventions. Inpatient hospital expenses and outpatient physician services do not change as much after adoption. There may be greater scope for increases associated with testing, imaging, operations, and inpatient pharmacy, since inpatient reimbursements are largely determined by the capitation-based Medicare diagnosis related group payments.

Usage of Diagnostic Testing

As described in the theory section, HIT is predicted to increase the reliance on diagnostic testing and imaging by increasing the intensity of treatment, but it is also predicted to reduce the rate of redundant testing. The net effect of these two impulses on diagnostic expenditures was found to be positive in the above section. In this section, I investigate the source of the increase in spending on diagnostics by analyzing electrocardiograms (EKGs). Results are reported in Table 4. EKGs are a very common form of diagnostic testing amongst the sampled patients, which measure the heart's rhythm and performance. The total number of EKGs per patient increases by 0.04 after HIT adoption, a result which is significant at the 5% level. Graphical evidence on the number of EKGs performed is presented in Figure 4, where EKG ordering appears to rise sharply following HIT adoption.

As reported in Table 4, column (3), the probability that a patient receives at least one EKG increased by about 0.6 percentage points, significant at the 10% level. Columns (5) and (6) look directly at the relationship between HIT adoption and redundant testing, analyzing the propensity for a patient to receive two or more EKGs within 30 days of their initial admission. By studying tests repeated within a relatively short window of time, I narrow the outcome to a measure more likely to include redundant, unnecessary testing. Although not all repeated tests are indications of redundancy, a decrease in this outcome variable would suggest that HIT helps providers reduce repetitive testing. HIT adoption is associated with a 0.6 percentage point increase in the likelihood of repeated testing, although the estimate is not statistically significant. The magnitude and statistical significance of the coefficient increases in the full trend-break model to a 0.9% increase in repeated testing three years post adoption, significant at the 5% level. Thus, HIT adoption is not

associated with a measurable reduction in redundant testing.

Research by Doyle et al. (2010) found physicians from a highly-ranked medical school spent less money on diagnostic testing per patient than physicians from a lower-ranked program, but achieved equivalent health outcomes. One explanation is that testing may substitute for the physician's cognitive time, medical knowledge, or critical reasoning abilities, which could otherwise be applied to diagnosing the patient. Providing HIT might further encourage this margin of substitution away from cognitive time towards increased testing by making it easier to order and follow test results. If there is a margin of substitution between testing and cognitive evaluation, shifting the margin may change costs without substantially affecting health outcomes.

Taken together, these results suggest that the propensity to order an EKG increased immediately after HIT adoption, and the increase came both from patients who became more likely to have a second, third, or subsequent EKG, as well as from patients who were more likely to receive a first EKG screen. By making tests easier to order and follow-up, HIT may encourage care providers to order more tests. Relatedly, physicians have told me in interviews that the value of a subsequent test is higher when they can easily view a prior test to track changes in the patient's status. These results suggest that a major predicted source of cost savings from HIT adoption, eliminating redundant testing, may not be realized.

Length of Stay and Reliance on Specialists

I now investigate the relationship between HIT and intensity of treatment by analyzing the impact on length of stay and reliance on medical specialists. Longer hospital stays signal a higher level of treatment intensity; results on length of stay

are reported in columns (1) and (2) of Table 5. HIT adoption is not associated with any substantial change in length of stay, with the 95% confidence interval bounding the effect between 1 hour shorter stay and 1 hour longer stay per patient, from a mean of 7 days per patient.

In addition, the total number of physicians seen within 1 year of admission does not change after HIT adoption. For this outcome, I separate the effects of clinical decision support (CDS) and electronic medical records (EMR) in the econometric model, since theory predicted that these technologies would have opposite effects on the number of specialists consulted. Results from this specification are reported in the bottom panel of Table 5. The point estimates are quite small in magnitude and not statistically significant. The coefficients have the opposite sign relative to the predictions of the theory, but the effects of EMR and CDS are not statistically distinguishable from each other, and each 95% confidence interval is bounded close to zero.

Since there is no evidence of significant heterogeneity of the effects of EMR and CDS, I combine both software types into the usual HIT adoption indicator to estimate the standardized composite effect. I find no significant relationship between HIT adoption and this measure of treatment intensity, as reported in Table 5, columns (5) and (6), as well as Figure 4. The 95% confidence interval around the estimate is bounded between a 0.01 standard deviation decline and 0.02 standard deviation increase in the intensity of treatment. This result confirms the finding that HIT adoption is not associated with economically substantial or statistically significant reductions in the costs and efficiency of care delivery, over the study period.

1.4.4 Impact of HIT on Hospital Quality

Quality of Inpatient Care

In this section, I analyze the impact of HIT on three measures of the quality of inpatient hospital care: 30-day readmission rate, complication rate, and adverse drug events. The results are reported in Table 6. Consistent with the null results on mortality, I find no impact of HIT on the 30-day readmission rate. A high readmission rate may indicate inadequate treatment of a patient's needs during their admission, and as such, poor quality of care. Incorrect prescriptions for the patient's home regimen and inadequate followup can also drive rising readmission rates. By improving the quality of inpatient care and making it easier to track the patient's medication list and construct an appropriate home regimen, HIT could reduce readmission rates. The 95% confidence interval bounds the coefficient between a 0.3 percentage point decline and a 0.06 percentage point increase in the readmission rate, from a mean of 8.4%. The lower bound suggests any reduction in readmission rate is not greater than 1 fewer readmission per 330 patients, from a mean of 28 readmissions per 330.

I similarly find no association between HIT adoption and reported complication rates, as reported in columns (3) and (4) of Table 6. Following Hougland et al. (2009), I measure the frequency of medical complications based on self-reported ICD-9 codes, which include errors (e.g. foreign object left in body, contaminated or infected blood transfusion) and complications (e.g. hemorrhage or infection due to procedure, abnormal reaction to surgery). The 95% confidence interval on medical complication rates bounds the estimate very close to zero: between a 0.4 percentage point reduction and 0.04 percentage point increase, from a mean of 6.5 percentage points.

Next, I analyze rates of adverse drug events. Rates of adverse drug events are also

constructed on the basis of self-reported ICD-9 codes, and include failures in dosage, accidental poisoning by drugs, or complications caused by the use of a medication (Hougland et al. 2009). This outcome is perhaps the one most directly linked to the specific features of the HIT software—medication lists, drug-drug interaction reminders, and drug allergy flags are all common components of popular HIT systems. In columns (5) and (6), I estimate slight increases in adverse drug events associated with HIT adoption. In column (5), the increase is a 0.14 percentage point increase, significant at the 10% level, which is equivalent to a 9% increase in the rate of adverse drug events. The effect is only marginally significant, but suggests that HIT adoption is *not* associated with reduced risk of pharmaceutical mismanagement.

Lastly, the standardized composite effect summarizes the findings across these three measures and finds no evidence of improvements in the quality of inpatient care. The composite effect is bounded between a -0.03 and 0.02 standard deviation change in the quality of care. Indeed, Figure 5 illustrates the flat path of the quality of care composite after HIT adoption.

Quality of Outpatient Preventive Care

Another dimension along which HIT has been predicted to improve the quality of care is by increasing adherence to preventive care guidelines (cf. Hillestad et al. 2005). EMR may increase the use of appropriate vaccines and screening services by making it easier for the clinician to observe whether the patient has recently received the service. CDS can provide reminders to physicians to order preventive care. Estimates from Bigelow et al. (2005) suggest that 52 million Americans over 50 have not received appropriate screening for colorectal cancer; and 19 million women over 40 have not had the recommended screening for breast cancer. It is thought

that there are large health benefits to receiving these early detection and preventive care services. Furthermore, if HIT can be shown to influence medical practice patterns for these widely accepted services, then it may hold promise for guiding physicians towards the recommended, evidence-based course of action, even as the specific guidelines and treatments evolve.

I study the impact of HIT adoption on the usage of colon cancer screenings and diagnostic mammograms billed by institutional providers for outpatient care. These results are reported in Table 7. An observation in these regressions is a hospital-year, and the outcome variable is the log count of screening services provided at that hospital. The sample is restricted to hospitals which bill at least one outpatient colon or breast cancer screening in every year from 1998-2005. Regression specifications for these outcome variables are similar to those described in Equations 1.8 and 1.9. Rather than controlling for the demographic characteristics of the inpatient population, I substitute a control for the size of the outpatient population serviced. Controlling for the number of outpatients seen at that hospital in the previous year ensures that an increase in the patient population is not confounded with higher rates of preventive care provision.

I find no significant effects of HIT on the provision of cancer screenings. The point estimates suggest modest increases in mammography screenings and small decreases in colon cancer screenings, but the effects are imprecisely measured and not statistically different from zero. In Table 7, the upper bound of the 95% confidence interval is a 5% increase in colon screenings and a 7% increase in mammograms. Increases of this magnitude would make slight progress towards bridging the large gap between the preventive care recommendations and the care that is actually delivered.

The standardized effect aggregating the two types of cancer screenings is reported in columns (5) and (6), as well as in Figure 6. I do no find evidence of a substantial

increase in the screening rates, aggregated across the two outcomes. The point estimate suggests 1% of a standard deviation increase in screening, with a confidence interval bounding a 0.02 standard deviation decline and a 0.04 standard deviation increase. In sum, I do not find strong evidence that HIT adoption is associated with increases in preventive care screenings, although modest effects cannot be ruled out.

1.4.5 Threats to Validity and Extensions

HIT Adoption and Patient Composition

The results suggest that HIT adoption is not associated with improvements in patient health, care quality, nor reductions in medical expenditures over the study period. One potential explanation for these findings is that after HIT adoption, the patient population being treated at the hospital becomes more complexly or acutely ill, along dimensions not fully captured by the included patient demographic and diagnosis controls. It could be that very ill patients select into hospitals that have adopted HIT systems, or that hospitals anticipate a change in their patient population and adopt HIT in response. These more ill patients would have worse health outcomes and higher medical expenditures, and could mask any improvements in care related to HIT adoption.

To address the concern of patient sorting, I test directly for compositional changes in the patient population after HIT adoption. Using patient characteristics as the outcome variables of interest, I run regressions analogous to the specifications in Equations 1.8 and 1.9. For these results, I omit patient and hospital characteristic controls from the right-hand side, in order to test directly for changes in these variables.

Results of these regressions are reported in Table 8, Panel A. In columns (1)

and (2), I find that HIT adoption is associated with slight 0.9% increases in the number of Medicare patients treated at the hospital. The increase is not statistically distinguishable from zero, but may suggest that HIT adopting hospitals are growing more quickly at the time of adoption.

However, even if adopting hospitals do serve a slightly growing patient population, I find no evidence that the patient disease or demographic characteristics change after HIT adoption. In columns (3) and (4), I test whether the diagnosis related group (DRG) weight associated with the patient's initial hospital admission increases after HIT adoption. Medicare calculates the DRG weight to facilitate its capitation based reimbursement for inpatient admissions. Reimbursement is a linear function of the DRG weight multiplied by a cost index and adjusted for geography. Hence, the DRG weight provides an indication of how complex the patient's medical needs are, since it is proportional to the usual costs of treating a patient with that particular constellation of diagnoses. Any increases in DRG weight could be evidence of either an increasingly complex and sick underlying patient population, or "up-coding" a patient's medical conditions enabled by the HIT software.

I find no evidence of changes to the diagnosis related group weight associated with HIT adoption. The point estimates are very small, a 0.005 increase in DRG weight, from a mean of 1.52, or less than a 0.3% increase in the estimated costs. The 95% confidence interval bounds the effect between a 0.006 decline in DRG weight and a 0.016 increase. The DRG weight does not indicate systematic changes in the illness of the underlying patient population.

I estimate a series of regression equations using patient demographic characteristics as the outcome variables of interest, including age, race, and sex. I then test the joint significance of the coefficients estimating the impact of HIT adoption on this set of characteristics across each equation, in a seemingly unrelated regression

framework. I cannot reject the null hypothesis that there is no change in patient demographics after HIT adoption; the p-value is 0.500 in the basic Equation 1.8 specification.

Similarly, I find no evidence of changes to the hospital's case-mix after HIT adoption. I test the joint significance of the coefficient on HIT adoption across a series of regression equations which use indicator variables for the patient's diagnosis as outcome variables. The p-value of this test is 0.967, indicating that I cannot reject the null hypothesis that there is no relationship between HIT adoption and case-mix.

The evidence presented in this section suggests that although adopting hospitals are growing at a faster rate than non-adopters, observable characteristics of the patient population are not changing abruptly right at the moment of HIT adoption. It is unlikely that changes in patient composition are driving the earlier null results on health outcomes and positive findings on expenditures and imaging frequency, although I cannot rule out the possibility that patient selection is changing along unobservable dimensions at the time of adoption.

HIT adoption and Other Hospital Investments

A second factor that may potentially confound estimates of the effect of HIT is that hospitals that choose to invest in HIT may be making other changes to their organizations. If hospitals were simultaneously investing in new imaging technology, for example, then these changes could be driving the estimated increases in diagnostic testing. In Table 8, Panel B, I analyze whether hospitals adopting HIT are also simultaneously investing in other costly medical technologies.

I construct an index of investment in medical technology, which is the sum of three indicator variables for whether the hospital has adopted a positron emission

tomography (PET) scanner, magnetic resonance imaging (MRI) scanner, and computed tomography (CT) scanner. Although HIT adopters are more likely to have each of these three technologies in the baseline year, investment in these technologies is uncorrelated with the timing of the HIT adoption decision. As reported in columns (1) and (2), HIT adoption is negatively related to the technology investment index, but the estimated coefficient is small and statistically insignificant. This finding alleviates concerns that the reported increases diagnostic imaging are driven by contemporaneous investments in imaging technology by adopting hospitals.

In columns (3) and (4), I analyze whether HIT adoption is related to a hospital's status as a trauma center. To obtain designation as a trauma center, a hospital must have the resources to evaluate and stabilize severely injured patients, including the capability for emergency resuscitation, surgery, and intensive care. Becoming a trauma center requires a hospital to maintain a specialized staff of physicians and surgeons large enough to provide uninterrupted emergency coverage, and may also require investment in a helicopter landing pad and other specialized equipment.

HIT adoption is negatively associated with the likelihood of being a designated trauma hospital. HIT adopters are 3 percentage points less likely to be designated trauma centers after adoption, significant at the 5% level. One explanation is that the large investments made in HIT adoption crowd out investment in becoming a designated trauma center. Trauma centers admit more critically ill patients, and thus may have worse health outcomes than other hospitals. The fact that HIT adopters are less likely to become trauma centers suggests that adopters may have a healthier patient population. This will bias the results towards finding a positive impact of HIT on health. To mitigate bias from this confounding factor, I directly control for trauma hospital status in all of the regression specifications reported in earlier tables, although adding this control does not materially change the results.

Lastly, I study whether hospitals adjust staffing inputs around the time of HIT adoption. Because the majority of hospitals do not directly employ their physician staff, the AHA survey data does not measure the number of physicians who are practicing and admitting to a particular hospital. The available measures of staffing include nursing and other support, facilities, and managerial staff. In columns (5) and (6), I find that HIT adoption is associated with modest, though statistically insignificant, 1.5% increases in staffing levels. This increase could be driven by the growing patient population or the increased need for IT support staff during the HIT implementation period.

In columns (7) and (8), I restrict the analysis to nursing staff. The number of full-time equivalent nurses increases by about 1.3% after HIT adoption, but the effect is imprecisely estimated and not distinguishable from zero. There is no evidence that HIT adoption is associated with reductions in the nursing staff, despite the fact that some of the functions automated by HIT systems may replace bookkeeping and report generating work previously done by nurses.

1.4.6 Robustness Tests

I further probe the robustness of my findings to the possibility of patient sorting outlined above. In particular, I aggregate the unit of observation from the hospital to the county level. If there is a greater scope for sorting across hospitals within a county, but little scope for sorting across counties, than these results will mitigate bias due to unobserved patient sorting.

These results are reported in Table 9, Panel A. The regressions estimated here echo the specifications in Equations 1.8 and 1.9, however instead of a binary variable for HIT adoption, the adoption variable now equals the fraction of patients within the

county treated at an adopting hospital. I omit the controls for differential trends by quartile of hospital size, since the observations are now aggregated to the county level. The point estimates are quite similar to those reported in earlier tables, although the standard errors are larger, making the findings less precise. Because aggregating to the county reduces precision without changing the overall findings, I prefer the hospital-level analysis for my main results.

In columns (1) and (2), I estimate a 1% increase in medical expenditures associated with HIT adoption, which is close to the 1.3% increase estimated in the hospital-level results. Similarly, the coefficient on patient mortality is extremely small in magnitude. Lastly, I find no economically large or statistically significant change in the intensity of treatment or quality of inpatient care standardized outcomes.

A second set of robustness checks is performed in Table 9, Panel B. Here, I test the concern that heterogeneity across adopting and non-adopting hospitals makes non-adopting hospitals a poor control group. Omitting the non-adopters from the regressions, I can estimate the impact of HIT using only adopting hospitals to estimate the time trends and impact of control variables.

The results omitting non-adopters are extremely close to the baseline regression estimates. HIT is associated with a 1.3% increase in medical expenditures and a 0.03 percentage point decline in the mortality rate, as before. The impact of HIT on the intensity and quality of care standardized outcomes remains small and not significant. The results presented earlier are not driven by the inclusion of never-adopting hospitals in the control group. In light of the consistent results, I prefer including non-adopters in the baseline specifications to improve the precision of the estimates.

1.4.7 Testing for Heterogeneous Returns to HIT Adoption

In the final set of results, I test whether the returns to HIT adoption are heterogenous either by the type of hospital adopting the software or the type of software being adopted. For the first set of results, reported in Table 10, Panel A, I test for heterogeneity in the impact of HIT by hospital size. One might predict that large hospitals have a larger ex ante coordination problem than small hospitals, so there could be further scope for HIT to benefit large hospitals. I compare hospitals in the top and bottom quartile based on the number of inpatient admissions reported in the AHA survey in the 1998 base year. Again, I run the regressions separately for each quartile and then test the equality of the coefficients across the two equations. Along this dimension, I do not find significant heterogeneity in the impact of HIT. Small hospitals experience larger increases in expenditures and larger declines in mortality, but the differences in the effect size between large and small hospitals are not statistically distinguishable.

Panel B tests whether hospitals with more comprehensive HIT installations experience greater returns to adoption. A leading hypothesis for why existing HIT systems appear to have modest, if any, impact on the quality and costs of care is that many of the current systems are not comprehensive enough to realize the full benefits to HIT adoption. Electronic medical records, clinical decision support, and a clinical data repository are thought to be three key ingredients for an effective HIT system (Fonkych and Taylor 2005). A clinical data repository integrates all electronic patient information into a single interface, and is thus thought to make the decision support and electronic record systems easier to use.

I test whether the effects of adopting a comprehensive HIT system, with all three software packages contracted, differ from the effects of the basic HIT components

studied in the previous sections. I code a binary variable indicating that a hospital has adopted a comprehensive HIT system if all three software packages are in place. Regressions follow the same structure outlined in Equations 1.8 and 1.9, but now each equation also includes the binary variable for adoption of a comprehensive HIT system. In addition, the trend-break model, based on Equation 1.9, includes an interaction between the comprehensive HIT adoption and normalized time.

I find no evidence that comprehensive HIT systems are associated with greater benefits to patient health, as compared to basic installations. Comprehensive HIT adoption is associated with more modest increases in expenditures than basic HIT systems, suggesting the full installation dampens the cost increases associated with the more limited installation. However, I cannot reject the null that the impact of the comprehensive system is equivalent to the impact of the basic system; the p-value of the test is 0.094.

Comprehensive HIT systems are associated with very small increases in mortality rates, where basic HIT systems are associated with a reduction in the mortality rate. However, both point estimates are very small and not statistically significant, and the difference between them is also not significant. The evidence presented in this section does not suggest that more comprehensive HIT systems will have substantially larger returns than the systems studied in the main results.[2]

1.5 Interpretation and Policy Predictions

On the whole, my results suggest that hospital HIT installations between 1998 and 2005 made little progress towards improving the quality and efficiency of the American healthcare system. HIT adoption was not associated with better health outcomes

[2]For a test of other alternative definitions of an HIT system, see results in Appendix Table A1.

or reduced costs. This runs contrary to the expectations of many policymakers and the optimism of the academic literature.

At the OLS estimated values of the coefficients, HIT adoption cost $1.9 million in additional spending for every additional life saved, not accounting for the installation costs of an HIT system. This is above the estimated $1.5 million value of remaining life for the average 75-year-old, as estimated by Murphy and Topel (2006), and above the $1.56 million value per statistical life estimated by Ashenfelter and Greenstone (2004). My in-sample patients are likely to have higher mortality rates than the average person of the same age, due to the selection on illness requiring inpatient admission, and so the Murphy & Topel value of a statistical life likely overstates the true value. Ashenfelter and Greenstone's estimate adjusts for neither patient age nor illness; thus it also provides an upper bound for the value relevant to the study population.

Accounting for the costs of installing an HIT system, the cost-benefit analysis becomes less attractive. I estimate the costs of installation and maintenance at $15 million over 10 years, using the more conservative estimates from the Southern California Evidence-based Practice Center (2006), or about $190 per inpatient admission. Including these expenses, an HIT system costs an estimated $2.5 million per life saved, at the estimated coefficient values.

One potential explanation for the low returns to HIT adoption is that unlike in other industries, where there were direct profit incentives to maximize the productive gains of a new technology, the incentives for physicians to use HIT as a tool to reduce billing and improve quality of care may be much weaker. The model presented in Section 2 of this paper demonstrated that HIT does not mechanically reduce billing costs, contrary to the conventional wisdom. Physicians have little incentive within most health care organizations and current reimbursement structures to reduce the

intensity of treatment they provide, so there is no a priori reason to believe that they will use HIT as a tool to accomplish this particular goal.

Bresnahan, Brynjolfsson, and Hitt (2002) find that it is often the combination of IT adoption and complementary organizational and technical innovation that leads to productivity gains. Cutler (2010) applies this argument to the healthcare sector, arguing that coupling information technology with organizational change may tremendously reduce the existing inefficiencies. Even if HIT could be a useful tool for reducing costs and raising the quality of care in a different institutional context, these returns may not be realized without changes to the current organizational structures and incentives facing care providers. Future research may explore this possibility by testing whether the returns to HIT vary across hospitals with different physician reimbursement and contracting structures. For example, care delivered at ambulatory surgical centers, where physicians may have ownership stakes and strong profit incentives to increase billing, could be contrasted with care delivered by managed care organizations or the Veterans Health Administration, where incentives are more likely to encourage reduced expenditures.

A second explanation for this finding is that HIT systems will evolve and improve, so future gains to adoption may greatly exceed the impact estimated here. The effects of HIT that I measure in this paper are local to the types of HIT commonly adopted over the study period. Within this data, I find no evidence that more comprehensive HIT systems are associated with substantially more favorable outcomes than basic IT installations. However, if new systems differ substantially from those currently offered, or if there are large positive synergies to adopting many ancillary components of an HIT system, then the results estimated here will not capture the full returns to HIT adoption.

A third possibility is that users take time to understand the software and learn

how to apply it effectively to their own work. This is a popular explanation for the productivity paradox of the 1980s and early 1990s, about which Robert Solow observed, "We see the computers everywhere but in the productivity statistics." David (1990) argues that information technology may require a long time scale to realize substantial returns, following the historical pattern of the diffusion and productivity gains of the dynamo. This possibility is difficult to test empirically, both because it requires a long data series, and because such a long-delayed relationship may be difficult to distinguish from other trending factors in the data. Nevertheless, the preceding analysis cannot rule out this possibility.

Within the limited window of the 7-year study period I analyze, the gains to HIT adoption do not appear to change substantially in the third, fourth, and fifth years after adoption. Indicators of patient mortality, readmissions, complications, and errors do not trend significantly downward several years after HIT adoption, as can be seen in Figures 2 and 5. The increase in total medical expenditures may be more transitory, with the regression results suggesting that the initial increases may not persist several years after adoption; however, there are no signs of substantial cost savings, even five years after initial adoption. The results do not suggest substantially higher returns to adoption even after physicians and nurses have had a few years to adjust to the new software systems.

A limitation of this analysis is that I do not test how HIT affects hospital operating costs; it may, for example, reduce bookkeeping costs or increase the number of outpatient clinic patients a physician can evaluate in a fixed amount of time. Lacking data on hospital's cost structures, this paper focused on measuring the health, quality of care, and medical billing impact.

The evidence presented here suggests that there is little social benefit to HIT adoption in the inpatient hospital setting. Patient health outcomes do not improve

significantly; expenditures by the Medicare social insurance program do not decrease. The argument for public subsidies of HIT adoption then hinges on the expectation that returns to HIT adoption will be higher in the future, perhaps in part due to innovation induced by the regulatory requirements of the HITECH act subsidies. Studying how the returns to HIT evolve remains a critical area of research as the HITECH legislation goes into effect over the next decade. A complementary research agenda would investigate whether particular organizational structures encourage higher-return adoption of HIT.

1.6 Conclusion

This study has analyzed the effects of Health Information Technology (HIT) adoption on the quality and intensity of medical treatment, both theoretically and empirically. It developed a model of inpatient healthcare delivery with attention to how the challenge of adapting care to a patient's idiosyncratic needs and coordinating care across disparate providers shapes treatment plans. I integrate physician effort, medical expenditures, and health outcomes into a single physician utility maximization model that may be useful for analyzing a variety of policy interventions. The model captures several key features of healthcare production which had not previously been unified into a single theoretical apparatus. The model provided key insight into the potential for HIT to increase the costs of care, a prediction that has not been widely discussed or acknowledged in the existing literature on HIT adoption.

The basis of the empirical analysis is a comparison of adopting hospitals before and after they first contract an HIT system. The impact of HIT adoption on Medicare patients receiving inpatient hospital care is measured using claims data from 1998-2005. Medical expenditures increase by approximately 1.3% after HIT adoption, in

particular due to increases in diagnostic testing and operative and pharmaceutical intervention. The cost increases are imprecisely measured though, and may not persist several years after adoption. The quality of hospital care, as measured by the mortality rate, readmissions, adverse drug events, and complications, does not change after HIT adoption. These results are robust to alternative specifications, including aggregating to the county level to mitigate potential bias due to patient selection, and omitting hospitals that never adopt HIT from the estimation sample. Overall, I find that HIT adoption is not associated with either cost savings or improved health outcomes over the study period. The evidence suggests that further research should be pursued into the conditions that might allow HIT to realize positive returns, before additional public money is spent.

1.7 References

Ashenfelter, Orley, and Michael Greenstone. 2004. "Using Mandated Speed Limits to Measure the Value of a Statistical Life." *Journal of Political Economy*, 112(S1): S226.

Aspden, Philip, Julie Wolcott, J. Lyle Bootman, and Paola D. Miralles. 2007. *Preventing Medication Errors: Quality Chasm Series.* Institute of Medicine.

Ayanian, John Z., Mary Beth Landrum, Edward Guadagnoli, and Peter Gaccione. 2002. "Specialty of Ambulatory Care Physicians and Mortality Among Elderly Patients After Myocardial Infarction." *New England Journal of Medicine*, 347(21): 1678–1686.

Baicker, Katherine, and Amitabh Chandra. 2004*a*. "Medicare Spending, the Physician Workforce, and Beneficiaries' Quality of Care." *Health Affairs*, 23: w18–w197.

Baicker, Katherine, and Amitabh Chandra. 2004*b*. "The Productivity of Physician Specialization: Evidence from the Medicare Program." *American Economic Review, Papers and Proceedings*, 94(2): 357–361.

Bates, David W, Deborah L Boyle, Eve Rittenberg, Gilad J Kuperman, Nell Ma'Luf, Valy Menkin, and Milenko J Tanasijevic MDD James W Winkelman MDD. 1998. "What Proportion of Common Diagnostic Tests Appear Redundant?" *The American Journal of Medicine*, 104(4): 361–368.

Bates, David W., Jonathan M. Teich, Joshua Lee, Diane Seger, Gilad Kuperman, Nell Ma'Luf, Deborah Boyle, and Lucian Leape. 1999. "The Impact of Computerized Physician Order Entry on Medication Error Prevention." *Journal of the American Medical Informatics Association*, 6(4): 313–321.

Becker, Gary S., and Kevin M. Murphy. 1992. "The Division of Labor, Coordination Costs, and Knowledge." *Quarterly Journal of Economics*, 107(4): 1137–1160.

Blumenthal, David, and John P. Glaser. 2007. "Information Technology Comes to Medicine." *New England Journal of Medicine*, 356(24): 2527–2534.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *The Quarterly Journal of Economics*, 117(1): 339–376.

Chandra, Amitabh, and Jonathan Skinner. 2003. "Geography and Racial Health Disparities." NBER Working Paper No. 9513.

Chaudhry, Basit, Jerome Wang, Shinyi Wu, Margaret Maglione, Walter Mojica, Elizabeth Roth, Sally C. Morton, and Paul G. Shekelle. 2006.

"Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care." *Annals of Internal Medicine*, 144(10): 742–752.

Congressional Budget Office. January 2008. "Technological Change and the Growth of Health Care Spending."

Cutler, David M. 2010. "Where Are The Health Care Entrepreneurs? The Failure of Organizational Innovation in Health Care." NBER Working Paper No. 16030.

Demakis, John G., Charles Beauchamp, William CUll, Robbin Denowwd, Seth A. Eisen, Richard Lofgren, Kristen Nichol, James Wooliscroft, and William G. Henderson. 2000. "Improving Residents' Compliance With Standards of Ambulatory Care: Results From the VA Cooperative Study on Computerized Reminders." *Journal of the American Medical Association*, 284(11): 1411–1416.

DesRoches, Catherine M., Eric G. Campbell, Christine Vogeli, Jie Zheng, Sowmya R. Rao, Alexandra E. Shields, Karen Donelan, Sara Rosenbaum, Steffanie J. Bristol, and Ashish K. Jha. 2010. "Electronic Health Records' Limited Successes Suggest More Targeted Uses." ***Health Affairs***, 29(4): 639–646.

Dessein, Wouter, and Tano Santos. 2006. "Adaptive Organizations." *Journal of Political Economy*, 114(5): 956–995.

Doyle, Joseph, Todd Wagner, and Steven Ewer. 2010. "Returns to Physician Human Capital: Evidence from Patients Randomized to Physician Teams." *Journal of Health Economics*, 29(6): 866–882.

Epstein, Ronald M. 1995. "Communication Between Primary Care Physicians and Consultants." *Archive of Family Medicine*, 4: 403–409.

Furukawa, Michael F., T. S. Raghu, and Benjamin B. M. Shao. 2010. "Electronic Medical Records and Cost Efficiency in Hospital Medical-Surgical Units." *Inquiry*, 47(2): 110–123.

Health Information and Management Systems Society. June 2, 2009. "Understand ARRA Spending Allocations with HIMSS Analytics Market Intelligence."

Hillestad, Richard, Anthony Bower James Bigelow, Federico Girosi, Robin Meili, Richard Scoville, and Roger Taylor. 2005. "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs." *Health Affairs*, 24(5): 1103–1117.

Himmelstein, David U., Adam Wright, and Steffie Woolhandler. 2010. "Hospital Computing and the Costs and Quality of Care: A National Study." *The American Journal of Medicine*, 123(1): 40–46.

Horwitz, Leora I., Thom Meredith, Jeremiah D. Schuur, Nidhi R. Shah, Raghavendra G. Kulkarni, and Grace Y. Jenq. 2009. "Dropping the Baton: A Qualitative Analysis of Failures During the Transition From Emergency Department to Inpatient Care." *Annals of Emergency Medicine*, 53(6): 701–710.e4.

Hougland, Paul, Jonathan Nebeker, Steve Pickard, Mark Van Tuinen, Carol Masheter, Susan Elder, Scott Williams, and Wu Xu. 2009. "Using ICD-9-CM Codes in Hospital Claims Data to Detect Adverse Events in Patient Safety Surveillance." Agency for Healthcare Research and Quality.

Javitt, Jonathan C., James B. Rebitzer, and Lonny Reisman. 2008. "Information Technology and Medical Missteps: Evidence from a Randomized Trial." *Journal of Health Economics*, 27: 585–602.

Jensen, Robert. 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *Quarterly Journal of Economics*, CXXII(3): 879–924.

Jha, Ashish K., Catherine M. DesRoches, Alexandra E. Shields, Paola D. Miralles, Jie Zheng, Sara Rosenbaum, and Eric G. Campbell. 2009*a*. "Evidence Of An Emerging Digital Divide Among Hospitals That Care For The Poor." *Health Affairs*, , (w1160).

Jha, Ashish K., Catherine M. DesRoches, Eric G. Campbell, Sara Rosenbaum, Karen Donelan, Timothy G. Ferris, Sowmya R. Rao, Alexandra Shields, and David Blumenthal. 2009*b*. "Use of Electronic Health Records in U.S. Hospitals." *New England Journal of Medicine*, 360(16): 1628–1638.

Jha, Ashish K., David C. Chan, Abigail B. Ridgway, Calvin Franz, and David W. Bates. 2009*c*. "Improving Safety And Eliminating Redundant Tests: Cutting Costs In U.S. Hospitals." *Health Affairs*, 28(5): 1475–1484.

Landon, Bruce E., Ira B. Wilson, Susan E. Cohn, Carl J. Fichtenbaum, Mitchell D. Wong, Neil S. Wenger, Samuel A. Bozzette, Martin F. Shapiro, and Paul D. Cleary. 2003. "Physician Specialization and Antiretroviral therapy for HIV." *Journal of General Internal Medicine*, 18(4): 233–241.

McCullough, Jeffrey S., Michelle Casey, Ira Moscovice, and Shailendra Prasad. 2010. "The Effect Of Health Information Technology On Quality In U.S. Hospitals." *Health Affairs*, 29(4): 647–654.

McKinsey Global Institute. 2002. "How IT Enables Productivity Growth."

Miller, Amalia, and Catherine Tucker. 2008. "Can Healthcare IT Save Babies?" Working Paper.

Murphy, Kevin, and Robert Topel. 2006. "The Value of Health and Longevity." *Journal of Political Economy*, 114(5): 871–904.

Murphy, Kevin, and Robert Topel, ed. 2003. *Measuring the Gains from Medical Research: An Economic Approach.* University of Chicago Press.

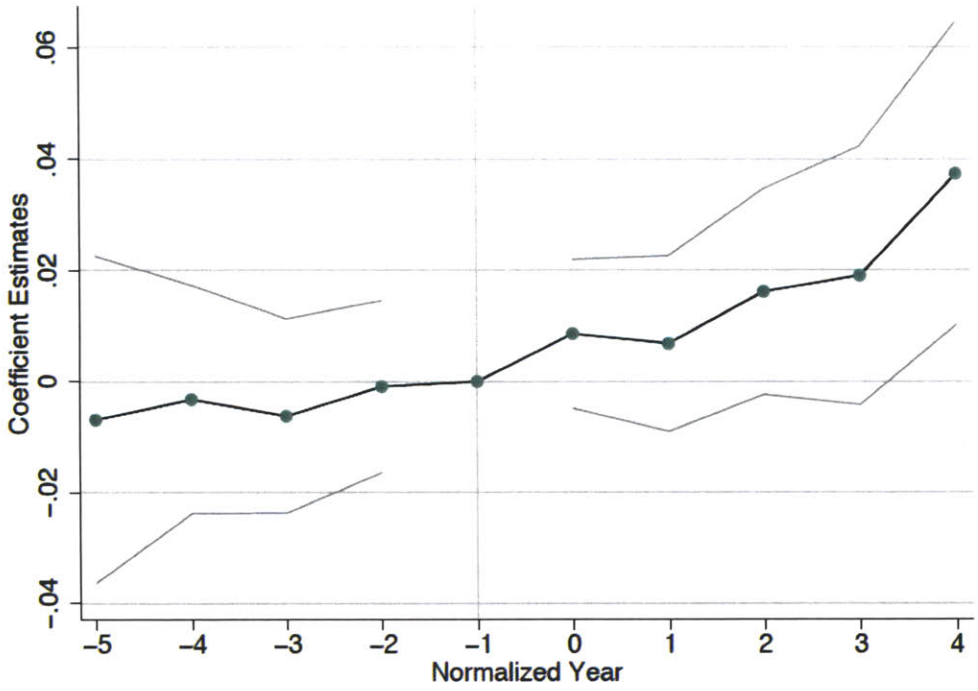
Poissant, Lise, Jennifer Pereira, Robyn Tamblyn, and Yuko Kawasumi. 2005. "The Impact of Electronic Health Records on Time Efficiency of Physicians and Nurses: A Systematic Review." *Journal of the American Medical Informatics Association*, 12(5): 505–516.

Southern California Evidence-based Practice Center. 2006. "Costs and Benefits of Health Information Technology."

Utah Office of Health Care Statistics. 2001. "Adverse Events Related to Medical Care Utah: 1995-99."

Wells, Kenneth B., and Roland Sturm. 1995. "Care for Depression in a Changing Environment." *Health Affairs*, 14(3): 78–89.

Figure 1: Log(1-Year Medical Expenditures)



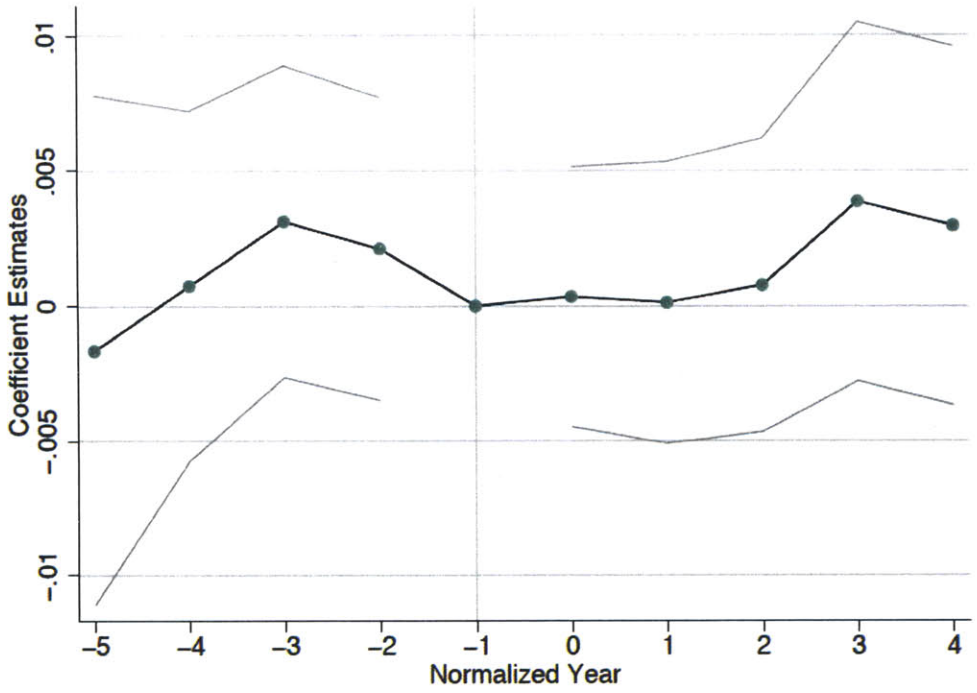
Notes:

The figure plots regression coefficients and 95% confidence intervals from a single regression where the dependent variable is Log(1-Year Medical Expenditures). The regression includes a series of explanatory dummy variables indicating the year relative to initial HIT adoption, for hospitals that change their adoption status over the study period. Adoption occurred in year 0.

Each regression controls for hospital fixed effects, state*year fixed effects, a differential time trend amongst adopting hospitals, as well as time trends that vary by quartile of hospital size. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis code. The 95% confidence interval is plotted in grey above and below the coefficient estimates. An observation is a hospital-year, 1998-2004. There are 27,317 observations in total.

Regressions are weighted by the number of patient observations that comprise the hospital-year observation, and standard errors are clustered by hospital.

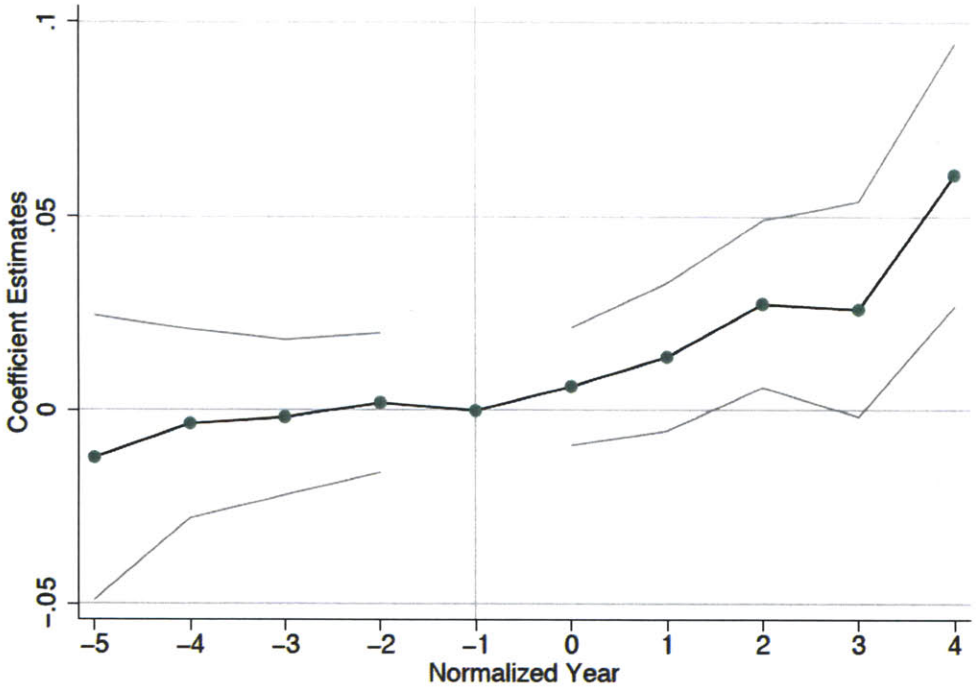
Figure 2: Mortality Rate



Notes:

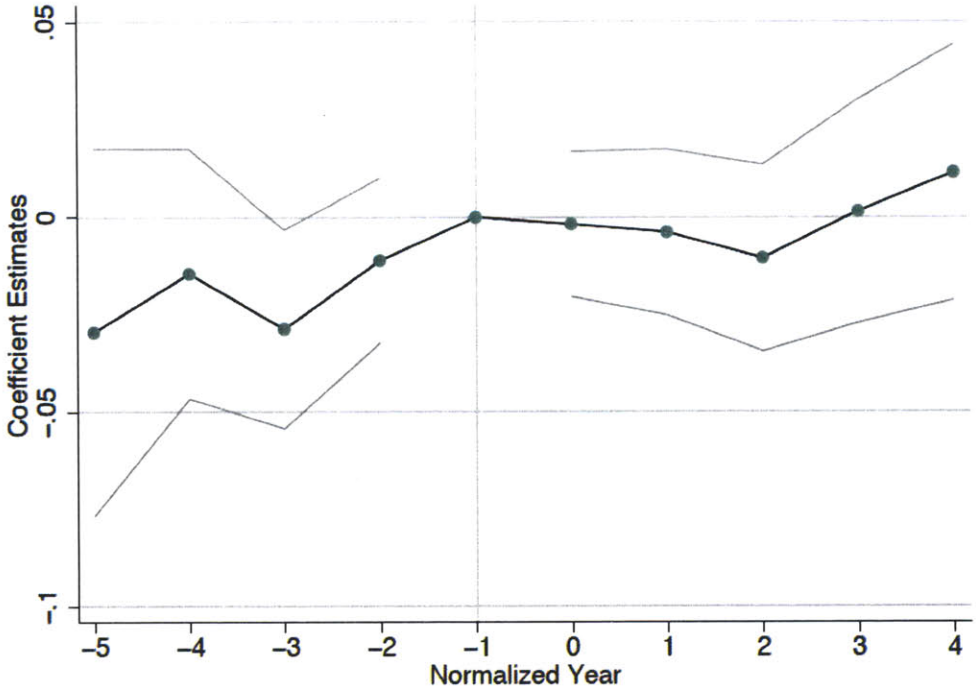
See notes to Figure 1. The figure plots regression coefficients and 95% confidence intervals from a single regression where the dependent variable is the 1-year mortality rate. The regression includes a series of explanatory dummy variables indicating the year relative to initial HIT adoption, for hospitals that change their adoption status over the study period. Adoption occurred in year 0.

Figure 3: Log(Expenditures on Diagnostic Testing & Imaging)



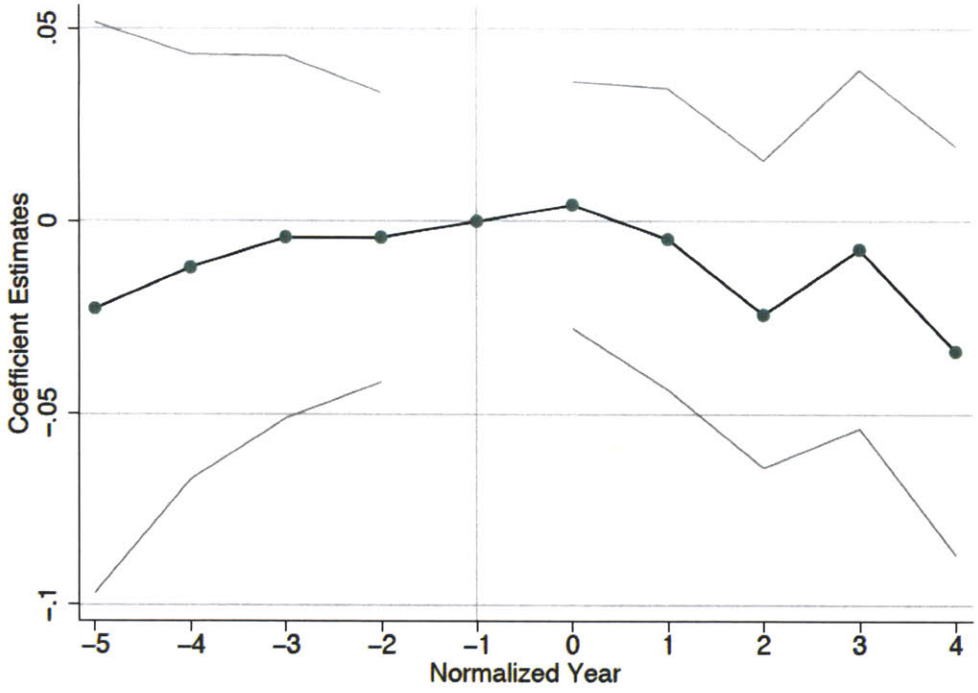
Notes:
See notes to Figure 1. The figure plots regression coefficients and 95% confidence intervals from a single regression where the dependent variable is the Log(Expenditure on Diagnostic Testing and Imaging). The regression includes a series of explanatory dummy variables indicating the year relative to initial HIT adoption, for hospitals that change their adoption status over the study period. Adoption occurred in year 0.

Figure 4: Intensity of Treatment Standardized Effect
Log(Length of Stay) & Log(Number of Physicians)



Notes:
See notes to Figure 1. The figure plots regression coefficients and 95% confidence intervals from a single regression where the dependent variable is a standardized effect that combines log(length of stay) and log(number of physicians seen) coefficients into a single index. The regression includes a series of explanatory dummy variables indicating the year relative to initial HIT adoption, for hospitals that change their adoption status over the study period. Adoption occurred in year 0.

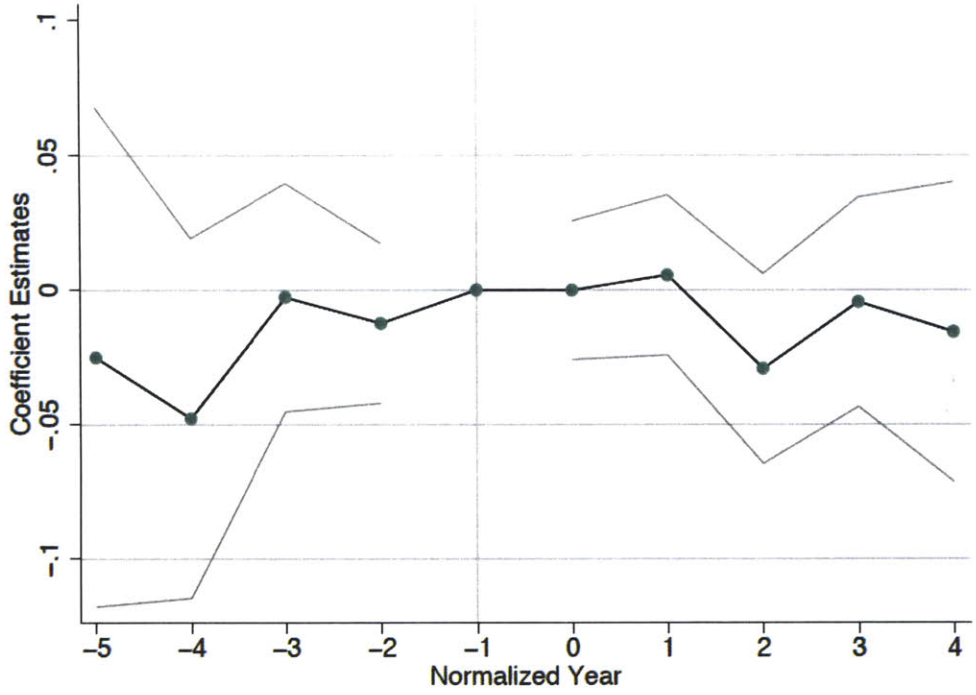
Figure 5: Quality of Inpatient Care Standardized Effect
Readmission, Complications, Adverse Drug Events



Notes:

See notes to Figure 1. The figure plots regression coefficients and 95% confidence intervals from a single regression where the dependent variable is a standardized effect that combines readmission rates, complication rates, and adverse drug event coefficients into a single index. The regression includes a series of explanatory dummy variables indicating the year relative to initial HIT adoption, for hospitals that change their adoption status over the study period. Adoption occurred in year 0.

Figure 6: Quality of Outpatient Preventative Care Standardized Effect Log(Screening Mammograms) & Log(Colon Cancer Screenings)



Notes:

See notes to Figure 1. The figure plots regression coefficients and 95% confidence intervals from a single regression where the dependent variable is a standardized effect that combines log(screening mammograms) and log(colon cancer screenings) coefficients into a single index. The regression includes a series of explanatory dummy variables indicating the year relative to initial HIT adoption, for hospitals that change their adoption status over the study period. Adoption occurred in year 0.

The regression controls for hospital fixed effects, state*year fixed effects, a differential time trend amongst adopting hospitals, as well as time trends that vary by quartile of hospital size. In addition, it controls for the number of unique patients seen at that hospital in an outpatient setting in the previous year.

Table 1: Summary Statistics, 1998 Baseline

	HIT Switchers (1)	Never HIT (2)	Always HIT (3)
A. Sample Size			
No. of hospitals	882	915	2086
No. of sampled patients per hospital	79	35	85
B. Hospital Characteristics			
No. of beds	209	101	223
Total admissions	8298	3300	9078
Total Medicare admissions	3187	1391	3487
FTEs	953	418	1075
Trauma hospital	0.300	0.235	0.322
Academic hospital	0.229	0.07	0.263
PET scanner	0.087	0.032	0.093
MRI machine	0.648	0.356	0.655
CT scanner	0.949	0.801	0.930
C. Sample Patient Characteristics			
Minority fraction	0.136	0.116	0.133
Age	77.0	77.8	76.6
D. Patient Outcomes			
Medical expenditures: 1 year	44,385	34,052	44,450
1-Year mortality	0.0993	0.106	0.099
E. Redundancy, Specialization, & Costs			
Average number of EKGs	1.4	1.15	1.5
Frac. receiving at least 1 EKG	0.222	0.151	0.226
Frac. receiving 2+ EKGs w/in 1 mo.	0.160	0.119	0.162
F. Intensity of Treatment			
Length of stay	6.9	6.6	7.0
Number of physicians: 1 year	13.3	10.5	13.4
G. Hospital Quality			
Medical complication	0.065	0.044	0.066
Medication error	0.016	0.019	0.018
Readmission	0.0843	0.082	0.085
H. Outpatient Preventative Care			
No. of screening mamograms	208	92	193
No. of colon cancer screenings	96	41	91

Notes:

All summary statistics are calculated on an annual basis for the 1998 base year. Sample includes hospital inpatients with one of the following diagnoses: acute myocardial infarction, stroke, pneumonia, hip fracture, lung cancer, colon cancer, or gastrointestinal bleed.

A hospital is considered an "HIT Switcher" if it adopts HIT between 1999-2004, and thus can be used to identify the effects of HIT adoption in the subsequent regressions. A hospital is in the "Always HIT" category if HIT was adopted prior to 1999. Hospitals in the "Never HIT" category have not adopted by the end of the study period.

Hospital characteristics are from the American Hopsital Association survey; HIT adoption is from the Health Information Management Systems Survey; all other variables are from the Medicare claims data. HIT is defined here as the adoption of at least one of the following technologies: Clinical Decision Support or Electronic Medical Records.

Table 2: Effect of HIT Adoption on Health and Total Expenditures

	Log(Medical Expenditures)		Patient Mortality	
	(1)	(2)	(3)	(4)
HIT Adoption	0.01285*	0.01761**	-0.00030	0.00055
	(0.00672)	(0.00828)	(0.00170)	(0.00193)
Post-Adoption Trend		-0.00328		-0.000588
		(0.00251)		(0.00074)
3-Year Effect		0.00778		-0.00121
		(0.00701)		(0.00212)
Mean of Dep. Var.	$44,385	$44,385	0.0993	0.0993

Notes:

The entries report the coefficients and standard errors (in parenthesis) from 4 separate regressions, where the dependent variable is Log(Medical Expenditures) in columns (1) and (2), and Patient Mortality in columns (3) and (4). Columns (1) and (3) report results from regressions that include HIT Adoption as the explanatory variable of interest. Columns (2) and (4) regressions include an HIT Adoption dummy and the interaction between time and HIT adoption.

Each regression controls for hospital fixed effects, state*year fixed effects, a differential time trend amongst adopting hospitals, as well as time trends that vary by quartile of hospital size. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis.

An observation is a hospital-year, 1998-2004. There are 27,317 observations. Regressions are weighted by the number of patient observations that make up the hospital-year observation and standard errors are clustered by hospital.

Rows denoted "3-Year Effect" report results of a test of significance of the linear combination of "[coef]HITAdoption + 3*[coef]interaction", which estimates the effect of HIT adoption 3 years after implementation.

HIT Adoption data from the Dorenfest Institute Survey; Hospital characteristics from the American Hosptial Association survey; patient demographics and outcomes from the Medicare claims data. *** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 3: Effect of HIT Adoption on Components of Medical Expenditures

	Log(Expend. on Diagnostics)		Log(Expend. on Inpatient Hospital Stay)		Log(Expend. on Pharmacy & Operations)		Log(Expend. on Physician Svcs, Outpatient Care)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HIT Adoption	0.01623*	.01955*	.01110	0.01520	.01512*	.02204**	-0.00929	-0.00854
	(0.00829)	(0.01026)	(0.00792)	(0.00981)	(0.00843)	(0.01074)	(0.01277)	(0.01597)
Post-Adoption Trend		-0.00228		-0.00282		-0.00476		-0.00052
		(0.00314)		(0.00288)		(0.00336)		(0.00527)
3-Year Effect		0.01270		0.00674		0.00776		-0.01009
		(0.00869)		(0.00807)		(0.00873)		(0.01394)
Mean Expend.	$9,887	$9,887	$14,324	$14,324	$15,027	$15,027	$5,148	$5,148

Notes:

See notes to Table 2. Entries are parameter estimates and clustered standard errors (in parentheses) from 8 separate regressions. The dependent variables are indicated in the column labels. Odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report results from regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption.

*** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 4: Effect of HIT Adoption on Diagnostic Testing

	Total Number of EKGs		Did the Patient Receive Any EKG?		Did the Patient Receive 2+ EKGs within 30 Days?	
	(1)	(2)	(3)	(4)	(5)	(6)
HIT Adoption	.03652**	.04704**	0.00557*	0.00506	.00568	0.00280
	(0.01775)	(0.02166)	(0.00335)	(0.00388)	(0.00366)	(0.00431)
Post-Adoption Trend		-0.00724		0.00036		0.00198
		(0.00704)		(0.00129)		(0.00127)
3-Year Effect		0.02531		0.00612		0.00875**
		(0.01948)		(0.00612)		(0.00394)
Mean Dep. Var.	1.4	1.4	0.22	0.22	0.16	0.16

Notes:

See notes to Table 2. Entries are parameter estimates and clustered standard errors (in parentheses) from 6 separate regressions. The dependent variables are indicated in the column labels. Odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report results from regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption.

*** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 5: Effect of HIT Adoption on Intensity and Efficiency of Care

	Log(Length of Stay)		Log(# of Physicians)		Standardized Intensity Composite	
	(1)	(2)	(3)	(4)	(5)	(6)
HIT Adoption	0.00087	-0.00195	0.00356	0.00743	0.00569	0.00382
	(0.00268)	(0.00334)	(0.00487)	(0.00670)	(0.00841)	(0.01103)
Post-Adoption Trend		0.00195		-0.00268		0.00130
		(0.00115)		(0.00210)		(0.00366)
3-Year Effect		0.00390		-0.00062		0.00771
		(0.00300)		(0.00461)		(0.00885)
Mean Dep. Var.	6.9	6.9	13.3	13.3		
			Separating Effects of EMR and CDS:			
CMS Adoption			0.002440	0.00557		
			(0.00598)	(0.00708)		
EMR Adoption			-0.00089	0.00191		
			(0.00389)	(0.00603)		
CDS Post Trend				-0.00185		
				(0.00149)		
EMR Post Trend				-0.00096		
				(0.00132)		
CDS=EMR?			0.675	0.899		
3 Year Comb. Effect			0.00155	-0.00097		
			(0.00623)	(0.00615)		

Notes:

See notes to Table 2. Entries are parameter estimates and clustered standard errors (in parentheses). The dependent variables are indicated in the column labels. In the first panel, odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report results from regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption.

Results reported in columns (5) and (6) combine estimates from the previous columns to estimate a standardized composite effect.

The bottom panel separates the effects of electronic medical records (EMR) and clinical decision support (CDS) by including separate explanatory variables for each type of software adoption. The "CDS=EMR?" row tests whether the estimated effect of clinical decision support equals the estimated effect of electronic medical records.

*** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 6: Effect of HIT Adoption on Quality of Inpatient Care

	30-Day Readmission Rate		Rate of Medical Complication		Rate of Adverse Drug Events		Standardized Inpatient Quality Composite	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HIT Adoption	-0.00126	-0.00102	-0.00181	0.00147	0.00143*	.00170*	-0.00332	0.00379
	(0.00097)	(0.00120)	(0.00114)	(0.00142)	(0.00074)	(0.00089)	(0.01440)	(0.01775)
Post-Adoption Trend		-0.00016		-0.00023		-0.00019		-0.00493
		(0.00041)		(0.00050)		(0.00032)		(0.00600)
3-Year Effect		-0.00152		-0.00218		0.00113		-0.01100
		(0.00109)		(0.00129)		(0.00088)		(0.01612)
Mean Dep. Var.	0.0843	0.0843	0.065	0.065	0.016	0.016		

Notes:

See notes to Table 2. Entries are parameter estimates and clustered standard errors (in parentheses). The dependent variables are indicated in the column labels. Odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report results from regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption.

Results reported in columns (7) and (8) combine estimates from the previous columns to estimate a standardized composite effect.

*** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 7: Effect of HIT Adoption on Outpatient Preventive Care

	Log(# of Screening Mammograms)		Log(# of Colon Cancer Screenings)		Standardized Preventive Care Composite	
	(1)	(2)	(3)	(4)	(5)	(6)
HIT Adoption	0.01809	0.03278	0.00868	-0.00890	0.01232	0.01085
	(0.02406)	(0.03501)	(0.02345)	(0.03730)	(0.01643)	(0.02448)
Post-Adoption Trend		-0.00646		0.00772		0.00064
		(0.00886)		(0.01109)		(0.00716)
3-Year Effect		0.01227		0.01330		0.01278
		(0.02135)		(0.00716)		(0.02184)
Mean Dep. Var.	208	208	96	96		

Notes:

The entries report the coefficients and standard errors (in parentheses), where the dependent variable is noted in the column labels. Odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption. Results reported in columns (5) and (6) combine estimates from previous columns into a standardized composite effect.

Each regression controls for hospital fixed effects, state*year fixed effects, a differential time trend amongst adopting hospitals, as well as time trends that vary by quartile of hospital size. In addition, it controls for the number of unique patients seen at that hospital in an outpatient setting in the previous year.

An observation is a hospital-year, from 1999-2005. There are 21,955 observations. Regressions are weighted by the number of patient observations that make up the hospital-year observation and standard errors are clustered by hospital.

Rows denoted "3-Year Effect" report results of a test of significance of the linear combination of "[coef]HITAdoption + 3*[coef]interaction", which estimates the effect of HIT adoption 3 years after implementation.

HIT Adoption data from the Dorenfest Institute Survey; Hospital characteristics from the AHA survey; patient demographics and outcomes from the Medicare claims data. *** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 8: Threats to Validity: Patient Selection & Hospital Investments

A. Patient Characteristics: Regression Results

	Log(# of Medicare Admissions)		Diagnosis Related Group Weight	
	(1)	(2)	(3)	(4)
HIT Adoption	0.00872	0.01162	0.00495	0.00326
	(0.01124)	(0.01258)	(0.00576)	(0.00691)
Post-Adoption Trend		-0.00200		0.00116
		(0.00394)		(0.00241)
3-Year Effect		0.00562		0.00675
		(0.01282)		(0.00666)
Mean Dep. Var.	3191	3191	1.524	1.524

B. Patient Characteristics: Omnibus Tests

	Patient Demographics: Age, Race, Sex		Patient Diagnosis Indicator Variables	
	(1)	(2)	(3)	(4)
p-value	0.500	0.417	0.967	0.994

C. Hospital Investments: Regression Results

	Technology Index		Trauma Hospital	
	(1)	(2)	(3)	(4)
HIT Adoption	-0.00626	-0.01202	-.032433**	-0.02849
	(0.02804)	(0.03492)	(0.01436)	(0.01937)
Post-Adoption Trend		0.00398		-0.00272
		(0.01100)		(0.00660)
3-Year Effect		-0.00009		-0.03666**
		(0.02973)		(0.01510)
Mean Dep. Var.	1.685	1.685	0.301	0.301

D. Staffing Inputs: Regression Results

	# of Full Time Employees		# of Full Time Nurses	
	(1)	(2)	(3)	(4)
HIT Adoption	0.01489	0.01771	0.01112	0.00733
	(0.00933)	(0.01123)	(0.01018)	(0.01265)
Post-Adoption Trend		-0.00194		0.00262
		(0.00328)		(0.00406)
3-Year Effect		0.01188		0.01519
		(0.00976)		(0.01096)
Mean Dep. Var.	954	954	240	240

Notes:

The entries in Panels A, C, and D report the coefficients and clustered standard errors (in parentheses) from 12 separate regressions, where the dependent variable is noted in the column labels for each panel. Odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption.

Panel B presents p-values from an omnibus test that tests the joint significance of the HIT adoption variable across equations. In columns (1) and (2), dependent variables are patient demographic characteristics. In columns (3) and (4), dependent variables are indicator variables for each in-sample diagnosis. Odd numbered columns directly test whether the coefficients on the binary HIT adoption variable are equal across all equations. Even numbered columns test the equality of the three-year effect.

Each regression (across all panels and columns) controls for hospital fixed effects, state-year fixed effects, a differential time trend amongst adopting hospitals, and time trends that vary by quartile of hospital size.

See notes to Table 2 for further details. *** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 9: Robustness Tests

	Log(Medical Expenditures)		Patient Mortality		Standardized Intensity Composite		Standardized Inpatient Quality Composite	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Aggregate Unit of Observation to County Level							
HIT Adoption	0.01004	0.00862	0.00009	-0.00011	0.01866	-0.00788	-0.01259	-0.00800
	(0.01040)	(0.01314)	(0.00240)	(0.00282)	(0.01453)	(0.01754)	(0.01830)	(0.02115)
Post-Adoption Trend		0.00099		0.00013		0.01447		-0.00250
		(0.00403)		(0.00118)		(0.00384)		(0.00455)
3-Year Effect		0.01159		0.00030		0.03552		-0.01551
		(0.01069)		(0.00315)		(0.01421)		(0.01835)
Mean Dep. Var.	$44,385	$44,385	0.0993	0.0993				
	B. Exclude Never-Adopting Hospitals							
HIT Adoption	0.01309*	0.01779**	-0.00025	0.00076	0.00703	0.00358	-0.00461	0.00072
	(0.00668)	(0.00827)	(0.00170)	(0.00194)	(0.00845)	(0.01109)	(0.01450)	(0.01787)
Post-Adoption Trend		-0.00323		-0.000692		0.00239		-0.00368
		(0.00254)		(0.00074)		(0.00368)		(0.00604)
3-Year Effect		0.00811		-0.00132		0.01074		-0.01033
		(0.00700)		(0.00212)		(0.00890)		(0.01622)
Mean Dep. Var.	$44,385	$44,385	0.0993	0.0993				

Notes:

Panels A & B: The entries report the coefficients and standard errors (in parentheses) from 16 separate regressions, where the dependent variable is noted in the column labels. Odd numbered columns report results from regressions that include HIT Adoption as the explanatory variable of interest. Even numbered columns report regressions that include an HIT Adoption dummy and the interaction between time and HIT adoption.

Panel A: An observation is a county-year, 1998-2004. There are 14,279 observations in total. Regressions are weighted by the number of patient observations that make up the county-year observation and standard errors are clustered by county. Each regression controls for county fixed effects, state*year fixed effects, and a differential time trend amongst adopting counties. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis.

Panel B: An observation is a hospital-year, 1998-2004. There are 21,068 observations in total. Regressions are weighted by the number of patient observations that make up the hospital-year observation and clustered by hospital. Each regression controls for hospital fixed effects, state*year fixed effects, and linear time trends that vary by quartile of hospital size. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis.

Panels A & B: Rows denoted "3-Year Effect" report results of a test of significance of the linear combination of "[coef]HITAdoption + 3*[coef]interaction", which estimates the effect of HIT adoption 3 years after implementation. The p-value of each test is reported immediately below each coefficient.

HIT Adoption data from the Dorenfest Institute Survey; Hospital characteristics from the AHA survey; patient demographics and outcomes from the Medicare claims data. *** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Table 10: Heterogeneous Returns to HIT Adoption

	Log(Total Expenditures)			1-Year Mortality		
	A. Hospital Size: Top vs. Bottom Quartile					
	Large Hospital	Small Hospital		Large Hospital	Small Hospital	
HIT Adoption	0.01452	.02888***		-0.00164	-0.00771	
	(0.01576)	(0.00389)		(0.00375)	(0.00173)	
p-value			0.376			0.141
	B. Type of HIT System: Basic vs. Comprehensive					
	Basic HIT System	Comprehensive HIT System		Basic HIT System	Comprehensive HIT System	
HIT Adoption	.01407**	0.00431		-0.00043	0.00055	
	(0.00680)	(0.00828)		(0.00171)	(0.00215)	
p-value			0.094			0.510

Notes:

Panel A: The entries report the coefficients and clustered standard errors (in parentheses) from 4 separate regressions, where the dependent variable is noted in the column header at the top. All regressions that include HIT Adoption as the explanatory variable of interest.

Regressions are run separately for the sub-sample of small hospitals, and the sub-sample of large hospitals. Each regression controls for hospital fixed effects, state-year fixed effects, and a differential time trend amongst adopting hospitals. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis. There are 5630 observations in the small hospital specifications and 6784 observations in the large hospital specifications.

Panel B: The entries report the coefficients and clustered standard errors (in parentheses) from 2 separate regressions, where the dependent variable is noted in the column header at the top. There are two explanatory variables of interest in each regression. A basic HIT system is defined as either clinical decision support or electronic medical record adoption. A comprehensive HIT system is defined as the adoption of both of those systems plus a clinical data repository.

Each regression controls for hospital fixed effects, state*year fixed effects, a differential time trend amongst adopting hospitals, as well as time trends that vary by quartile of hospital size. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis. There are 27,317 observations.

Both Panels: The "p-value" row reports the p-value from a test of equality of the two coefficients listed in the preceding columns. Regressions are weighted by the number of patient observations that make up the hospital-year observation.

HIT Adoption data from the Dorenfest Institute Survey; Hospital characteristics from the AHA survey; patient demographics and outcomes from the Medicare claims data. *** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Appendix Table A1: Analysis of Computerized Order Entry and Other Ancillary Systems

	Log(Medical Expenditures)		Patient Mortality		Standardized Intensity Composite		Standardized Inpatient Quality Composite	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			A. Impact of Computerized Physician Order Entry					
CPOE Adoption	0.03054***	0.03278***	0.00343	0.00444	0.00637	0.00666	0.01341	0.01799
	(0.00971)	(0.01128)	(0.00266)	(0.00284)	(0.02041)	(0.02434)	(0.01467)	(0.01657)
Post-Adoption Trend		-0.00314		-0.00142		-0.00042		-0.00650
		(0.00596)		(0.00171)		(0.01384)		(0.00828)
3-Year Effect		0.02337		0.00019		0.00540		-0.00149
		(0.01508)		(0.00492)		(0.03410)		(0.02207)
Mean Dep. Var.	$44,385	$44,385	0.0993	0.0993				
	B. Impact of Computerized Physician Order Entry, Laboratory Info. Sys, & Radiology Info. Sys.							
CPOE, LIS, RIS	0.01555**	0.01518*	0.00188	0.00222	0.00847	0.00879	-0.01403	-0.01433
	(0.00732)	(0.00791)	(0.00193)	(0.00200)	(0.01586)	(0.01719)	(0.01056)	0.01119
Post-Adoption Trend		0.00098		-0.00090		-0.00085		0.00079
		(0.00546)		(0.00166)		(0.01263)		(0.00782)
3-Year Effect		0.01811		-0.00047		0.00624		-0.01195
		(0.01507)		(0.00486)		(0.03465)		(0.02229)
Mean Dep. Var.	$44,385	$44,385	0.0993	0.0993				

Notes:

See notes to Table 2. Entries are parameter estimates and clustered standard errors (in parentheses) from 16 separate regressions. The dependent variables are indicated in the column labels. For these results, the usual explanatory variable is replaced with an indicator variable for the adoption of computerized physician order entry (CPOE) in Panel A or the simultaneous adoption of CPOE, a laboratory information system, and a radiology information system in Panel B.

Odd numbered columns report results from regressions that include technology adoption as the explanatory variable of interest. Even numbered columns report results from regressions that include an HIT Adoption dummy and the interaction between time

Each regression controls for hospital fixed effects, state*year fixed effects, a differential time trend amongst CPOE adopting hospitals, as well as time trends that vary by quartile of hospital size. Additional controls include patient age (in 1-year bins), sex, race, and primary diagnosis.

An observation is a hospital-year, 1998-2004. There are 27,317 observations. Regressions are weighted by the number of patient observations that make up the hospital-year observation and standard errors are clustered by hospital.

Rows denoted "3-Year Effect" report results of a test of significance of the linear combination of "[coef]HITAdoption + 3*[coef]interaction", which estimates the effect of HIT adoption 3 years after implementation.

HIT Adoption data from the Dorenfest Institute Survey; Hospital characteristics from the AHA survey; patient demographics and outcomes from the Medicare claims data. *** denotes significance at 1% level; ** denotes significance at 5% level; * denotes significance at 10% level.

Discussion:

As with the main results on eletronic records and cliical decision support, computerized physician order entry systems (CPOE) are associated with increased costs, although the increase may not be permanent. Mortality rate and quality of care measures do not improve after HIT adoption.

Results are similar for hospitals adopting CPOE alongside laboratory information systems and radiology information systems. Expeditures rise after adoption, although more modestly, while there are no signs of improvment in patient health or quality of care.

Chapter 2

Diffusion of New Medical Technologies: Evidence on Physician Learning [1]

2.1 Introduction

The adoption of new medical technologies has driven substantial gains in longevity as well as steep expenditure growth in the health care industry over recent decades. Improving the process by which physicians learn about new technologies and incorporate them into their medical practice is thus critical to realizing value in the health care industry. In this paper I investigate the determinants of early adoption of new

[1]I would like to thank David Autor, Amy Finkelstein, and Michael Greenstone for invaluable guidance throughout this project. I thank Jason Abaluck, Joshua Aronson, Joseph Doyle, Jonathan Gruber, Danielle Li, Amanda Pallais, Heidi Williams, and participants at MIT's labor lunch for their comments. I am also grateful to Mohan Ramanujan and Jean Roth for their assistance in obtaining and managing the data. John Wang provided excellent research assistance. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship.

technologies, as well as the scope for physician learning across geographic regions. In particular, I analyze whether late adopting regions are able to effectively apply the lessons of early adopters when determining how to use a new diagnostic or therapeutic technology. I find evidence that late adopters do not fully repeat the process of experimental learning undertaken by early adopters, suggesting that there may be efficiency gains associated with staggering the adoption of new medical technologies.

The Congressional Budget Office (2008) has estimated that about half of the 4.9% real rate of growth in health spending has been driven by technology adoption. Some of this technology adoption has had tremendously high returns. For example, the survival probability of low birthweight babies has improved by 12 percentage points with a rate of return of over 500% (Cutler and Meara 2000). On the other hand, some high cost technologies, such as angioplasty, appear to be overused despite low marginal returns, while other technologies with low costs and high estimated health benefits, such as beta blockers, appear underused. These inefficiencies, against the backdrop of rising costs, suggests that a richer understanding of what drives technology adoption could be critical to formulating policies that would ultimately improve the efficiency with which new technologies are adopted or discarded.

This study focuses on the diffusion of two medical technologies: positron emission tomography (PET) scans and deep brain stimulation. PET scanning is a nuclear imaging technique that produces a three dimensional image of cellular metabolic activity. The technology is now widely used in clinical oncology, and in more limited use to diagnose certain neurologic diseases and to map heart function. PET scanning was first approved in 1995 for analyzing heart function, and the indications for usage were progressively expanded between 1998-2005. Over the study period, it was most frequently applied to patients with lung cancer. PET scanners require large capital investments of $1 million to $2.5 million, as well as learning about the relevant

clinical applications and scan interpretation required to implement the technology effectively.

The second technology is deep brain stimulation, a surgical treatment for neurological movement disorders. The procedure involves the implantation of a device that sends electrical impulses to targeted areas of the brain, suppressing involuntary movement. The procedure was FDA approved in 1997 for essential tremor, in 2002 for Parkinson's Disease, and in 2003 for dystonia. The most frequent application over the study period is to Parkinson's Disease. In addition to these explicitly covered diagnoses, physicians have also experimented with deep brain stimulation as a treatment for epilepsy and, in rare cases, psychiatric conditions. As a novel intervention for advanced neurological disease, it is targeted to a narrow population and requires significant new surgical training to adopt the technology.

Using Medicare claims data, I demonstrate that there is substantial heterogeneity in the timing of technology adoption across states. For PET scans, the extent of the market is correlated with early adoption and application of the scanning technologies, likely driven by the high fixed costs of investment. This correlation is not present for deep brain stimulation, perhaps because the economies of scale are smaller for this intervention.

I find that the mix of patients treated with the new technologies changes substantially during the early stages of diffusion, as physicians learn about which patients benefit from the intervention. Notably, late adopting states do not appear to repeat the learning experience of early adopters; rather, their patterns of technology usage mirror the contemporaneous usage patterns of early adopters, albeit with less population penetration. The medical returns to technology usage do not evolve measurably over the study period for these two technologies.

The paper proceeds as follows. Section 2 provides further discussion of the context

and data, comparing the characteristics of early and late adopters. Section 3 outlines the empirical strategy and the highlights how the applications of the technologies have evolved over time. Section 4 presents results related to the changing returns to technology usage. Section 5 concludes by analyzing the policy implications and limitations of these findings.

2.2 Industry Context & Data

2.2.1 Technology diffusion and physician learning

The adoption of new medical technologies is often associated with high initial costs. As with PET scanners, a large capital outlay may be required. Clinical staff must be trained on the application and operation of the technology. Physicians will need to learn which patients can benefit from the intervention, and what the potential costs, side effects, and treatment synergies may be. In addition, clinicians may need time to perfect their technique: developing appropriate dosages, treatment schedules, or honing surgical skill.

Understanding the determinants of technology adoption and the relative importance of peer learning versus learning-by-doing mechanisms is critical to formulating optimal social policy and reimbursement coverage for new medical interventions. In particular, policymakers must grapple with whether reimbursement permission should be staggered across geographic regions, whether early adopters should be subsidized, and the optimal timing of widespread insurance reimbursement allowance. In the US context, these policy issues are particularly pressing given that Medicare exercises significant market power, providing insurance for one in seven Americans, and private industry frequently follows Medicare's lead in approving coverage of new

medical technologies (Van de Water 2008).

This research builds on the recent literature on technology diffusion and learning in the healthcare industry. Several earlier studies have analyzed the determinants of medical technology adoption (Coleman et al. 1966, Taub et al. 2011, Azoulay 2002, Berndt et al. 2003), and found that new scientific evidence, physician age, degree of specialization, social networks, and market concentration all play a role in determining a physician's decision to adopt a new technology. Other work by Skinner and Staiger (2005) documented that there is significant geographic correlation between the adoption of effective medical technologies and the state-level patterns of adoption of hybrid corn, tractors, and computers. They propose that state-level human capital may be a critical explanatory variable behind these patterns of diffusion. This body of work describes who chooses to adopt a new technology, but does not directly test whether these patterns are socially efficient, and if there are gains from the observed staggering of adoption.

Recent work by Ramanarayanan (2008) and Huckman and Pisano (2006) find evidence of learning-by-doing and firm-specific learning, respectively, amongst cardiac surgeons. This research describes some of the ways in which physician learning may be specific to his own experience and practice context, but does not directly test how effectively learning about new technologies can be transferred across physicians. By exploiting a national panel data set, this paper provides insight into how the patterns of technology usage evolve amongst early and late adopters, demonstrating how provider learning manifests itself in the patterns of technology usage.

2.2.2 Medicare claims data

The primary data source for this analysis is a 20% sample of Medicare Part A and Part B claims, for a period of eight years from 1998 until 2005. These include claims rendered for inpatient hospital care and for physician services delivered in both an inpatient and outpatient context. Using this data, I construct a national panel data set on deep brain stimulation and PET scan usage, permitting comparisons across regions, over time. I observe the diagnoses, comorbidities, and prior medical expenditures of patients selected to receive treatment with the new technology, and contrast these characteristics with those of the broader population of patients with related diagnoses. In addition, data on patient mortality and health outcomes allow an analysis of the evolving returns to technology usage.

One limitation of this data is that I can only observe technology adoption and usage amongst the population of Medicare patients. If physicians tend first to apply a new technology to non-Medicare patients, then there may be measurement error in the identification of early- and late-adopting regions. In addition, estimates of the learning costs associated with technology adoption would be biased towards zero, if much of the costs are borne by patients not included in this sample.

These concerns are mitigated by the fact that the diseases targeted by deep brain stimulation and PET scanning are much more prevalent amongst the elderly, Medicare-eligible population. For Parkinson's Disease, the most frequent indication for deep brain stimulation, the average age of onset is 60, and deep brain stimulation is a treatment option only for late stages of the disease (Reider et al.; National Institute of Health 2010). The most common recipients of PET scans over this period were lung cancer patients, 68% of whom are diagnosed after the age of 65 (National Cancer Institute). Due to the high concentration of these diseases amongst

the elderly, it is less likely that a physician would adopt the technology exclusively for the non-Medicare population.

The analysis contrasts the adoption experience at the state level, comparing diffusion of the new technologies across early- and late-adopting states. I focus on the state-level adoption experience for a few reasons. Foremost, finer levels of geographic gradation make it very difficult to identify precisely which areas are early adopters due to the small number of in-sample patients receiving treatment with the new technology. I have implemented this analysis at the county-level, and the results are qualitatively similar, but much less precise. Secondly, medical licensure, statewide professional organizations, hospital regulation, and exposure to malpractice law, all operate at the state-level, and thus, the state is a natural unit of analysis for understanding variation in medical practices.

States are grouped into early- and late-adopting regions according to the earliest year in which technology usage exceeded the 25th percentile of technology penetration as a fraction of the target population for that technology in that state. For PET scanning, this corresponds to the year that 0.1% of patients with eligible diagnoses received at least one scan; thirty-one states reached this threshold by 2000 and are categorized as early adopters. For deep brain stimulation, the threshold corresponds to 0.07% of patients with eligible diagnoses receiving treatment; 27 states reached this threshold by 1999 and are categorized as early adopters. The results that follow are not sensitive to these precise definitions and are robust to alternative thresholds for the definition of "early" and "late" adopters. Figures 1 and 2 present maps illustrating which states are early adopters for each technology.

The usage of PET scanners expanded rapidly over the study period, with the average penetration amongst eligible patients increasing 30-fold from 0.02% in 1998 to 0.6% in 2005. Figure 3 contrasts the diffusion of PET scanners across early and

late adopting states, plotting the fraction of patients with eligible diagnoses receiving a scan over time. The 95% confidence interval around each point estimate is plotted in grey. The difference in the penetration rates between early and late adopters increases through the early years of the diffusion process; at the peak in 2002, early adopters are performing 0.2 percentage points more scans per eligible patient than late adopters, or 88% more scans.

Figure 4 displays the diffusion pattern for deep brain stimulation. Due to the relative rarity of this intervention, the diffusion curves are estimated with less precision, and distinguishing early from late adopters becomes more difficult. Compared to the case of PET scanning, the initial differences between the early and late adopters in 1998 are more substantial, suggesting that amongst early adopters, many of them took up the technology immediately following FDA approval in 1997. The gap between the early and late adopters closes over the study period, and the levels of adoption are not statistically distinguishable in the later years. However, an F-test rejects the joint hypothesis that early and late adopters have equivalent diffusion levels in each year.

2.2.3 Summary statistics: comparing early and late adopters

Table 1 contrasts the characteristics of early and late adopters in the baseline year of 1998. In Panel A, it is reported that early adopters of PET scans have an 85% greater population of patients with eligible diagnoses, compared to late adopters. PET scanning is a technology that likely exhibits high returns to scale, since there are high fixed costs to purchasing and staffing the machine which can then be used for many patients.

In addition to the greater population of eligible patients in early adopting states,

the mortality rate of patients with lung cancer diagnoses is substantially higher in early adopting states, and the patients receive more intensive medical treatment with greater spending and more days spent as a hospital inpatient. Thus, early adopting states not only have more eligible patients, but also a greater population of severely ill patients.

In Table 1, Panel B, we see that early adopters of deep brain stimulation actually have smaller populations of patients diagnosed with Parkinson's Disease or essential tremor. Since the marginal costs of providing each additional surgery are likely very high relative to the fixed costs of offering the service, the size of the market may be less critical to the adoption decision.

Patients with Parkinson's Disease appear to be less severely ill and receive less costly medical intervention in early adopting states relative to the late adopters. Since the success of an intervention designed to mitigate the symptoms of Parkinson's Disease is unlikely to have a direct impact on mortality, I use an alternative health outcome for these patients: the incidence of hip fractures. Tremor symptoms exacerbate the risk of falling and inhibit the patient's ability to break his fall, and as a result, patients with poorly controlled Parkinson's Disease are more likely to have hip fractures (Pressley 2003). The rate of hip fractures amongst all patients with a diagnosis of Parkinson's is 1.7% in early adopting states, compared to 1.9% in late adopting states.

A limitation of the state-level analysis that follows is that there is substantial heterogeneity in the baseline rates of disease prevalence and severity across regions. The observed patterns of adoption are endogenous to to the state's demand for the technology, including patient characteristics and physician proclivity to adopt new interventions. The context is not a natural experiment and the self-selection of state-level technology adoption must be considered when interpreting the results. In

particular, to the extent that the medical returns and patients receiving treatment differ across early- and late-adopters, some of these differences may be attributable to differences in the patient population or physician skill and preferences, rather than directly attributable to a learning mechanism or lack thereof. The concern is even more important in light of the significant unobserved patient heterogeneity in this context; because I cannot identify a clinically equivalent patient seeking treatment in two different states, I rely on aggregate statistics to make comparisons about the propensity to offer the new intervention and the returns to its application. These limitations are discussed further below, in the context of the results.

2.3 Applications of New Medical Technologies

As physicians experiment with a new medical technology, they learn both about the value of the intervention and about which patients benefit most from treatment. This learning process shapes how many patients receive treatment with the new technology, which types of patients receive the treatment, and what the medical returns are to these interventions. I explore each of these three outcomes of the learning process in turn.

2.3.1 Do late adopters have accelerated diffusion paths?

First, I analyze how the number of patients receiving treatment evolves over time. If physicians find a technology to be clinically or financially valuable to their practice relative to alternative treatment modalities, usage of the technology should increase over the stages of diffusion as more physicians adopt the technology and adopting physicians increase their intensity of usage. In this way, the demonstrated success

of technology adoption in one area may accelerate the adoption in another area. Amongst late adopters, physicians who are just beginning to use the technology will have less uncertainty about the technology's value, and so will apply the technology more intensively and more rapidly to match the diffusion pattern in other regions.

Figure 5 plots the coefficients and 95% confidence interval from a regression that provides year-by-year estimates of the level of PET diffusion. Time has been normalized so that year 0 is the first year in which the state has achieved at least the 25th percentile level of diffusion or 0.1%. The normalized trends are displayed for early adopters who reached this threshold by 2000, and late adopters who reached the threshold in 2001 or later.

Figure 5 displays little evidence of accelerated adoption amongst later adopters of PET scanners. The diffusion curves appear to be nearly parallel, with the late adopters a little over one year behind the early adopters throughout the entire study period. Regression analysis bears this out: an F-test fails to reject the hypothesis that the normalized year indicators differ across early and late adopters. The late adopters do not appear to be catching up to earlier adopters over this period.

Analogous results for deep brain stimulation are presented in Figure 6. Here, time has been normalized so that year 0 is the first year in which the state has achieved at least a 0.07% adoption level; early adopters have reached this threshold by 1999. The picture is murkier for deep brain stimulation, as compared to PET, for a few reasons: the earliest stages of adoption are unobserved for the first wave of adopters; and the small sample of patients receiving treatment reduces the precision of the diffusion estimates. Again, an F-test fails to reject the hypothesis that the normalized year coefficients are equal for early- and late-adopters.

There are a few potential explanations for this finding that the new technologies did not diffuse more quickly amongst late adopters. First, returns to technology

adoption may be state-specific, so late adopters would not learn very much about the technology from early adopters. The argument that returns are highly heterogeneous across states is weakened by the fact that late adopters appear to be expanding their use of the technology at a similar rate as early adopters, and there is no evidence within the frame of this sample that they settle on a significantly lower steady-state adoption level. However, to fully test for differences in the steady state level of technology usage, a longer time series would be required.

A second explanation is that resource or capacity limitations may dictate the speed of diffusion within a region, and learning about a technology's returns may have limited influence on the speed with which it is adopted. Lastly, it is possible that information about the value of a technology or the conditions under which the technology has positive returns is difficult to communicate or does not diffuse perfectly. In this case, even if the returns to adoption are similar across states, late adopters may not have accelerated adoption paths. To disentangle these potential explanations, I turn to the evolving patterns of which patients are targeted with a new technology.

2.3.2 Do late adopters apply new technologies for the "right" diagnoses?

Once a technology has received FDA approval, physicians have significant leeway to apply the technology more broadly than the approved indications. Moreover, the initial indications for treatment may not correspond to the patients with the greatest medical return. With both PET scans and deep brain stimulation, there were substantial changes in the mix of patients receiving treatment over the observed stages of adoption.

One test of the scope for information transfer across regions is whether late adopting states repeat that experimentation process undertaken by the early adopters. In particular, does the case mix of a late adopter in its first or second year of technology usage mirror the first or second year of technology usage by an early adopter? Or do late adopters immediately adopt the current best practices of the more experienced regions? If there are risks or costs associated with the early stages of adoption, it would be more beneficial to stagger adoption as long as the late adopters do not need to repeat the complete learning experience of earlier adopters.

To test whether late adopting states repeat the experimentation of early adopters, applying a new technology to lower return patients, I estimate regressions where the outcome variable is the share of scans or surgeries performed for a particular indication in state s at time t. The regression takes the following form:

$$\begin{aligned} \text{TechnologyShare}_{st} = {} & \alpha_1 + \alpha_2 \text{LateAdopter}_s \text{EligibleShare}_{st} + \beta_1 \text{Yr}_t \text{EligibleShare}_{st} + \\ & \beta_2 \text{Yr}_t \text{LateAdopter}_s \text{EligibleShare}_{st} + \gamma \text{EligibleShare}_{st} + \epsilon_{st} \end{aligned} \quad (2.1)$$

LateAdopter_s is an indicator variable for whether the observation is from a late adopting state. Yr_t is a running variable for the current calendar year. $\text{EligibleShare}_{st}$ is a variable indicating the fraction of the potentially eligible population with the diagnosis of interest. Each included variable is interacted with the fraction eligible to capture the fact that states with smaller eligible populations will mechanically have those diagnoses represent a smaller fraction of treated patients, and that this smaller population will attenuate any potential growth in treatment for those indications. I also report results from an alternative un-interacted linear specification, but the interacted model is the preferred specification for the aforementioned reason. Standard errors are clustered at the state level.

If there are limited cross-region or nationally centralized learning mechanisms, then late adopters must repeat some of the learning experience of early adopters. In particular, we would expect α_2 to have the opposite sign of β_1, indicating that the late adopters are some years behind the early adopters in their choice of clinical applications, after controlling for any differences in their patient population. In the extreme case of no learning, we would additionally expect that β_2 would be zero, i.e. that the slope of their learning curve will be identical to that of the earlier adopters. With imperfect information transfer, we might expect β_2 would have the same sign as β_1, indicating an accelerated trend for late adopters, as they catch up to the current best practices. Perfect information transfer would be characterized by α_2 and β_2 both equal to zero, i.e. the late adopters exactly match the patterns of clinical application chosen by the earlier adopters at a given point in time.

In addition to the specification parameterized above, I also non-parametrically estimate the year-by-year changes in the technology's applications separately for early and late adopters. This is accomplished with a regression that includes a series of indicator variables for the year of the observation, and an interaction between the current year and whether the state is a late adopter. Plotting these coefficients ensures that nonlinear time trends are not creating misleading results in the linear regression framework outlined above.

In Table 2, Panel A, the evolving applications of PET scans are documented. First approved for mapping of heart function in 1995, PET became increasingly applied to cancer patients, rather than heart disease patients, over time. Although the total number of patients provided with PET scans increased across all eligible diagnoses with greater technology availability and familiarity, the growth was much more rapid in cancer diagnoses than heart disease diagnoses. In 1998, 47% of PET scans performed were for cancer-related indications; by 2005, 71% of PET scans were

for cancer patients.

Regressions reported in Table 2, Panel A, estimate that amongst early adopting states, there were 18% per year more patients with eligible cancer diagnoses receiving scans. In the non-interacted, additive specification this amounts to a an additional 3.4 percentage points of total performed PET scans were applied for cancer diagnoses rather than for non-cancer indications each year, a result that is statistically significant at the 1% level. Amongst late adopters, a larger fraction of PET scans were performed for cancer indications from the beginning, although the difference is not statistically significant, and there is no evidence that the rate of growth in cancer scans differed for late adopters. The point estimate on the interaction between time and late adoption is small in magnitude at -0.6 percentage points per year, and not statistically distinguishable from zero. These results are illustrated in Figures 7a and 7b. Figure 7a illustrates the growing share of cancer-related PET scans performed by early adopting hospitals. Figure 7b plots the difference in the cancer share of PET scans amongst later adopters relative to early adopter's benchmark in a given year; the coefficient estimates hover near zero for the duration of the period.

As reported in Figures 8a and 8b, the inverse pattern emerges in the analysis of heart scans as a share of total scans, with heart scans declining over the study period. Again, there is no evidence that late adopters repeat the learning process undertaken by early adopters.

A competing explanation for these findings, besides efficient transfer of medical learning across states, is that late adopting hospitals have more patients suitable for PET imaging with cancer diagnoses, relative to early adopters. In this case, the fact that their treatment patterns match those of the early adopting hospitals may be due to differences in their patient population, and not due to transferred learning about which patients may benefit from the technology. The share of patients with the

relevant eligible diagnosis is an imperfect proxy for the proportion of the population suitable to the new treatment, so this control may not fully capture differences in the patient population. However, along all observable dimensions, the population of late adopting states is, if anything, less suited to cancer applications on average, compared to early adopters. 18.9% of eligible patients in early adopting states have cancer diagnoses in 1998, compared to only 18.1% of eligible patients in late adopting states. Moreover, the early adopting states have a more ill population of cancer patients, with a higher mortality rate, which would be associated with indications for more frequent and intensive monitoring with PET technology.

Within the set of PET scans performed on cancer patients, there were also changes in the composition of patients targeted. In particular, a growing fraction of cancer scans were performed on lymphoma patients relative to lung cancer patients. Again, the late adopters appear to copy the current practice of early adopters in selecting which cancer patients should receive the new treatment. Neither the interaction term nor the level shift is statistically distinguishable from zero. The nonparametric regression results are plotted in Figures 9 and 10.

In the case of deep brain stimulation, applications to Parkinson's Disease, the most common indication, made up a relatively constant share of total cases. However, essential tremor cases became a shrinking share as applications to convulsions, epilepsy, and other less common indications grew. The evidence for deep brain stimulation in Table 2, Panel B, and in Figures 11 and 12, illustrates a similar pattern as the evidence for PET scanning. Late adopters do not appear to repeat the learning undertaken by early adopters about which indications should be receiving treatment with the new technology. The intercept shift and interaction term on the time trend are not statistically distinguishable from zero in any of the specifications. Although late adopters do not experience accelerated diffusion curves relative to early adopters,

they do appear to learn from early adopters about the appropriate indications for treatment.

2.3.3 Do late adopters apply technologies to the "right" types of patients?

In addition to analyzing the changing diagnosis mix of treated patients, I also study prior patterns of medical resource utilization to analyze how the disease severity of treated patients evolves throughout the technology's diffusion path. I analyze three measures of disease severity, all constructed using the patient's inpatient medical claims from the year preceding their treatment with the new technology: days spent over previous year as a hospital inpatient, prior year's inpatient hospital charges, and the Charlson comorbidity index. The comorbidity index is constructed as the weighted sum of a patient's chronic health conditions and is designed to predict a patient's mortality risk. In this context, it provides an indication of how precarious the patient's medical condition was, particularly as it related to secondary diagnoses.

For this analysis, I focus on the most common indication for each technology: lung cancer patients receiving PET scans, and Parkinson's Disease patients receiving deep brain stimulation. The regression framework is similar to that outlined above in equation 1, and it takes the following form:

$$\begin{aligned} \text{PatientHealthStatus}_{nst} = {} & \alpha_1 + \alpha_2 \text{LateAdopter}_s \text{Treated}_n + \beta_1 \text{Yr}_t \text{Treated}_n + \\ & \beta_2 \text{Yr}_t \text{Treated}_n \text{LateAdopter}_s + \gamma_t + \delta_t \text{LateAdopter}_s \end{aligned} \quad (2.2)$$

The outcome variable is an indicator of the patients' health status for patients in state s in year t with treatment status n, where n indicates whether the patient received

treatment with the new technology. These regressions include patients untreated with the new technology so that time trends in medical resource utilization and comorbidity reporting can be allowed to flexibly differ across early and late adopting states. The inclusion of these patients identifies a set of year fixed effects that differ according to the state's status as an early or late adopter, γ_t and δ_t. The coefficient β_1 on the interaction $Yr_t Treated_n$ summarizes how the set of patients receiving treatment with the new technology evolves within early adopting states. The coefficients α_2 and β_2 describe the difference between late adopters' and early adopters' treated patients in a given year.

To improve the power of the tests and reduce the rate of false positive results, I also create standardized effect measures across these outcomes. The standardized effect is calculated separately for lung cancer applications of PET scans and for Parkinson's applications of deep brain stimulation, and includes the hospital charges, days as an inpatient, and comorbidity index calculated over the previous year. These groupings allow me to perform omnibus tests analyzing whether selection into treatment with the new technology trends towards relatively healthier or more sick patients within a particular diagnosis, over time. I report both separate results for each outcome variable, as well as the aggregated standardized effect. I account for the cross-equation covariance structure of the error terms when estimating standard errors for each outcome within a domain. Standard errors remain clustered at the state level.

The standardized effect is constructed by combining the estimated coefficients across each outcome variable within a domain. In particular, the standardized effect equals:

$$\sum_{j \in J} \frac{1}{J} \frac{\beta_{ij}}{\sigma_j}, , i \in \{1, 2\} \tag{2.3}$$

where β_{1j} is estimated by equation 2 for outcome variable j, or similarly for α_2 and β_{2j}. σ_j is the standard deviation of the outcome j, in the baseline year of 1998. Dividing by the standard deviation harmonizes the units across the diverse outcome variables. J is the total number of outcomes within a domain.

The evidence presented in Table 3, Panel A suggests that over time, PET scanning was applied to increasingly sick lung cancer patients, i.e. patients with higher hospital charges, more inpatient days, and higher comorbidity rates, in the year before receiving a PET scan. Each year, patients receiving treatment were 0.035 standard deviations less healthy as measured by the standardized effect, which is statistically significant at the 5% level. The coefficient estimate α_2 suggests that patients receiving scans in the late adopting states are slightly healthier than those in the early adopting states, putting the late adopters approximately one year behind the early adopters in their application of the technology. However, this difference between late and early adopters is imprecisely estimated and not statistically distinguishable from zero. In addition, the time trend in these patient selection variables for late adopters is almost identical to the trend for early adopters. Table 3, Panel B, presents an analogous set of results for patients receiving deep brain stimulation, although the coefficient estimates are too imprecise to infer the patterns of diffusion for these outcome variables.

Taken together, the results presented in Table 3 neither contradict nor corroborate the pattern that emerged in the previous section, with later technology adopters immediately applying the current practices of the more experienced adopters. Trends in these patient characteristics are smaller in magnitude and less precisely estimated than the trends in diagnosis mix, but are not inconsistent with the possibility of learning.

2.4 Medical Returns to Technology Adoption

As physicians learn about the applications and value of a new technology, the medical returns to technology usage may evolve. The direction of the trend on medical returns is theoretically ambiguous. If the first group of treated patients happens to be those with the highest medical returns, physicians may choose to move down the curve and expand treatment to the next most suitable patient group, lowering the average return. On the other hand, for a given patient receiving treatment, the treatment may become better executed over time, as the physician learns about the optimal delivery system for the new technology.

I test empirically how the medical returns to technology adoption evolve, contrasting the experience of early and late adopters. The regression framework parallels that listed in equation 2 above, replacing patients' pre-treatment characteristics with post-treatment outcomes as the dependent variables. The outcomes employed are hospital charges, days spent as an inpatient, mortality, and hip fracture rate, all calculated over one year following treatment with the new technology. As above, I include controls for year fixed effects that differ according to a state's adoption timing, using untreated patients to identify the trends. I also calculate standardized effects to summarize the health outcome measures into a single index.

A challenge in estimating the health returns to technology adoption is that selection into treatment is changing over time, and it is difficult to disentangle the selection effect from the evolving returns to treatment on a given patient. The problem is particularly challenging given the coarse clinical diagnosis and health status variables available in the claims data. In an effort to address this concern, I report results from a second set of regressions, which are augmented with additional controls for the pre-period health status of treated patients, i.e. charges, inpatient days,

and comorbidity index. These controls are included to distinguish changes in the medical returns to technology usage from changes in the selection into the treatment group. Contrasting regression results with and without these controls provides some insight into the drivers of any changes in the health returns to technology usage.

Results from this analysis are reported in Table 4. Estimated trends in the standardized health outcome composite measure are not statistically distinguishable from zero; I cannot reject the hypothesis that there is no change in the medical benefits over the course of the diffusion path. This null result holds up across both the lung cancer population receiving PET scans and the Parkinson's Disease patients receiving deep brain stimulation, and is not sensitive to the inclusion of controls for the treated patient's pre-period health status. It should be noted, however, that the estimates are sufficiently imprecise that I cannot rule out economically meaningful changes in a patient's mortality rate or medical expenditures. Moreover, some of the point estimates on individual outcomes are statistically significant at conventional levels and suggest that there may be some limited trending.

Notably, the mortality rate of treated patients is increasing for early adopters by 1.7 percentage points per year, after controlling for observable health status, significant at the 5% level. It is unlikely that PET scans are driving such a large increase in patient mortality, which further highlights the difficulty of accurately controlling for patient health status. Excluding controls for health status, the mortality rate is declining over time amongst PET recipients in early adopting regions. The pre-period health status of treated patients is also declining, and it would appear that these proxies for health status understate the true decline in the pre-period health of treated patients, driving an increase in the adjusted mortality rate. To accurately capture the medical returns to technology adoption for a particular type of patient, richer clinical data would be invaluable.

2.5 Conclusion

Taken as a whole, the evidence presented above suggests that there is some scope for the valuable transfer of information about a technology's applicability across states in the US. In particular, it appears that late adopting states do not repeat the process of experimentation already undertaken by early adopters to learn which diagnoses benefit most from treatment with the new technology. There are a number of potential mechanisms for this information transfer, from formal journal articles and conference presentations, to informative advertising by medical device manufacturers and informal information sharing across physician's social networks. Disentangling these channels would be a fruitful line for future research, shedding light on effective ways of new disseminating information about medical best practices.

A major limitation of the preceding analysis is the absence of exogenous variation in the timing of technology adoption. Because the adoption decision is a choice variable, it is likely to be correlated with the patients' suitability and physicians' preferences for the technology. As a result, it is possible that the observed patterns of the evolving diagnosis mix across early and late adopting regions are driven not by learning but by fixed characteristics of the states. Although there is nominally regional variation in reimbursement policy for Medicare, Agha (2011) finds that the local coverage determinations have no measurable influence on actual services rendered, so the Medicare program rules provide little traction. Future work could make use of the timing and location of clinical trials or variation in the generosity of private insurance, for example, to provide cleaner variation in the timing of adoption.

Understanding how the returns to the application of new technology evolve is critical to making normative evaluations of the learning process. Unfortunately, claims data is ill-suited to identifying the health benefits associated with a new technology,

both because measures of health outcomes are limited, and because patient heterogeneity is not very finely observed, and thus difficult to control for. There could be considerable economic insight gained from the analysis of richer chart data, which may allow much better measurement of the returns to technology adoption.

The construction of optimal reimbursement policy for new technologies depends on understanding both the health and financial costs associated with early experimentation as well as the scope for learning across physicians and hospitals. In this paper, I find evidence that physicians do learn from the adoption experience of peers in other states, and apply those lessons when they begin using a new technology. Information about the appropriate diagnoses to be targeted by a new intervention is costly to discover but appears easy to communicate, and late adopters of deep brain stimulation and PET scanners apply the technology to the same mix of patients as earlier adopters in a given year.

2.6 References

Agha, Leila. 2011. "Managing Medical Technology Policy: A Review of Current Practices and Future Directions."

Azoulay, Pierre. 2002. "Do Pharmaceutical Sales Respond to Scientific Evidence?" *Journal of Economics and Management Strategy*, 11(4): 551–594.

Berndt, Ernst R., Robert S. Pindyck, and Pierre Azoulay. 2003. "Network Effects and Diffusion in Pharmaceutical Markets: Antiulcer Drugs." *Journal of Industrial Economics*, 51: 243–270.

Coleman, James S., Elihu Katz, and Herbert Menzel. 1966. *Medical Innovation: A Diffusion Study.* Bobbs-Merrill Company.

Congressional Budget Office. January 2008. "Technological Change and the Growth of Health Care Spending."

Cutler, David M., and Ellen Meara. 2000. *The Technology of Birth: Is It Worth It?* Vol. 3, MIT Press.

Huckman, Robert S., and Gary P. Pisano. 2006. "The Firm Specificity of Individual Performance: Evidence from Cardiac Surgery." *Management Science*, 52(4): 473–488.

National Cancer Institute. 2011. "Surveillance Epidemiology and End Results Cancer Statistics Review 1975-2008."

National Institute of Neurological Disorders and Stroke, National Institute of Health. n.d.. "NINDS Deep Brain Stimulation for Parkinson's Disease."

Pressley, J.C., E.D. Louis, M.-X. Tang, L. Cote, P.D. Cohen, S. Glied, and R. Mayeux. 2003. "The impact of comorbid disease and injuries on resource use and expenditures in parkinsonism." *Neurology*, 60: 87–93.

Ramanarayanan, Subramaniam. 2008. "Does Practice Make Perfect: An Empirical Analysis of Learning-by-Doing in Cardiac Surgery." Working Paper.

Reider, Carson R., Cheryl A. Halter, Peter F. Castelluccio, David Oakes, William C. Nichols, Tatiana Foroud, and the Parkinson Study Group. 2003. "Reliability of Reported Age at Onset for Parkinson's Disease." *Movement Disorders*, 18(3): 275–279.

Skinner, Jonathan, and Douglas Staiger. 2005. "Technology Adoption from Hybrid Corn to Beta Blockers." NBER Working Paper No. 11251.

Taub, Anna A. Levine, Anton Kolotilin, Robert S. Gibbons, and Ernst R. Berndt. 2011. "The Diversity of Concentrated Prescribing Behavior: An Application to Antipsychotics." NBER Working Paper No. 16823.

Van de Water, Paul N. 2008. "Medicare Change Can Complement Health Reform." Center on Budget and Policy Priorities.

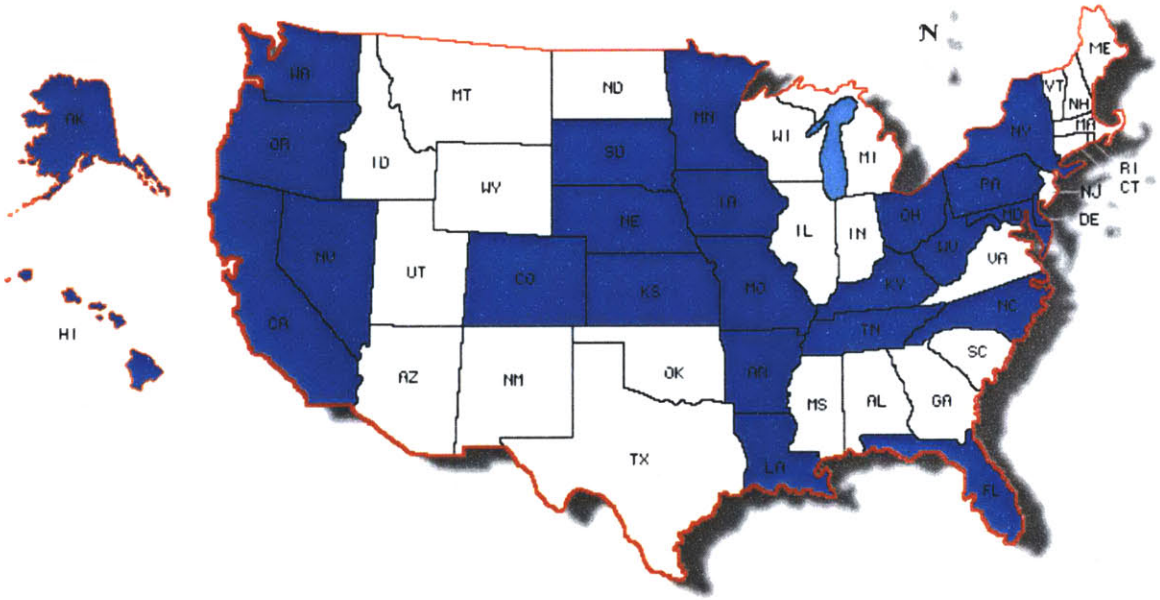


Figure 1: Map of PET Adoption.

Notes: Early adopters are shaded in blue. A state is considered an early adopter of PET scans if 0.1% of eligible patients received a scan by 2000.

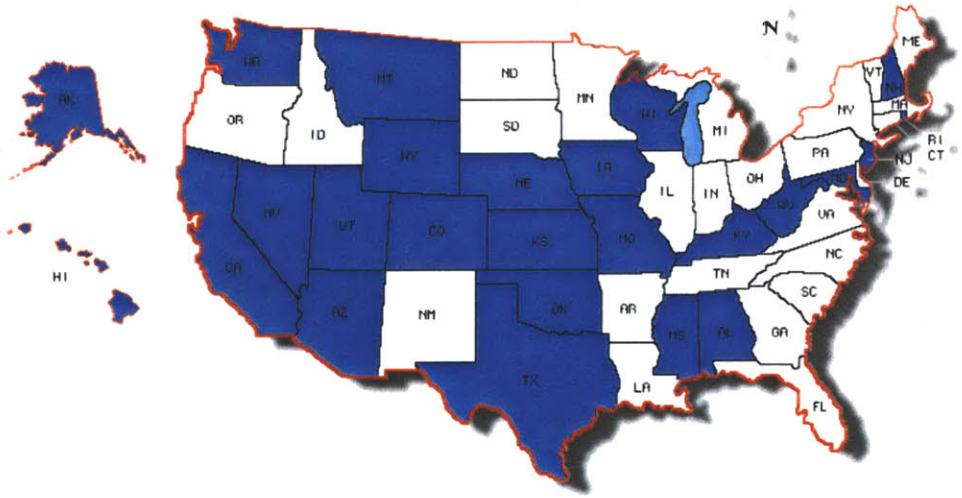


Figure 2: Map of Deep Brain Stimulation Adoption.

Notes: Early adopters are shaded in blue. A state is considered an early adopter of deep brain stimulation if 0.07% of eligible patients received surgery by 1999.

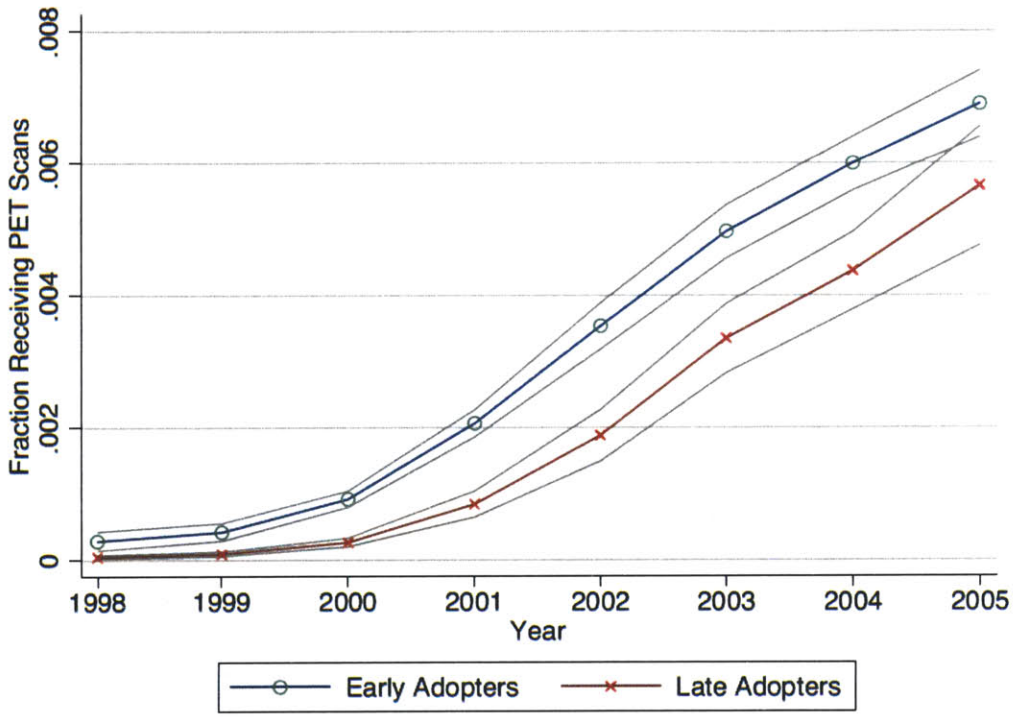


Figure 3: Diffusion of PET Scans

Notes: This is a plot of state-level diffusion of PET scans as a fraction of the total number of eligible patients. The 95% confidence interval is plotted in grey.

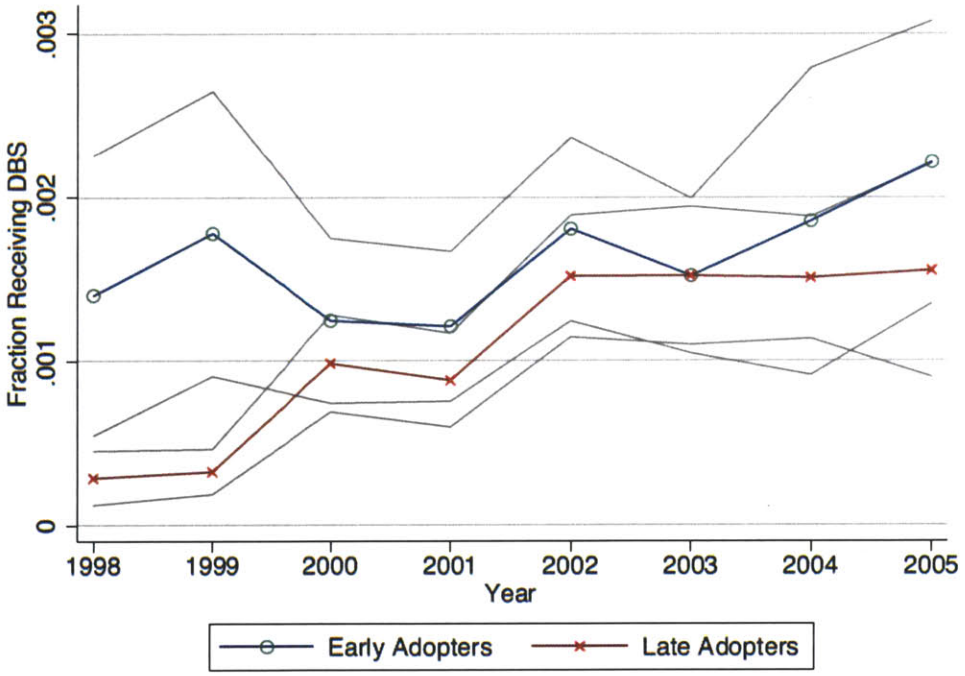


Figure 4: Diffusion of Deep Brain Stimulation

Notes: This is a plot of state-level diffusion of deep brain stimulation as a fraction of the total number of eligible patients. The 95% confidence interval is plotted in grey.

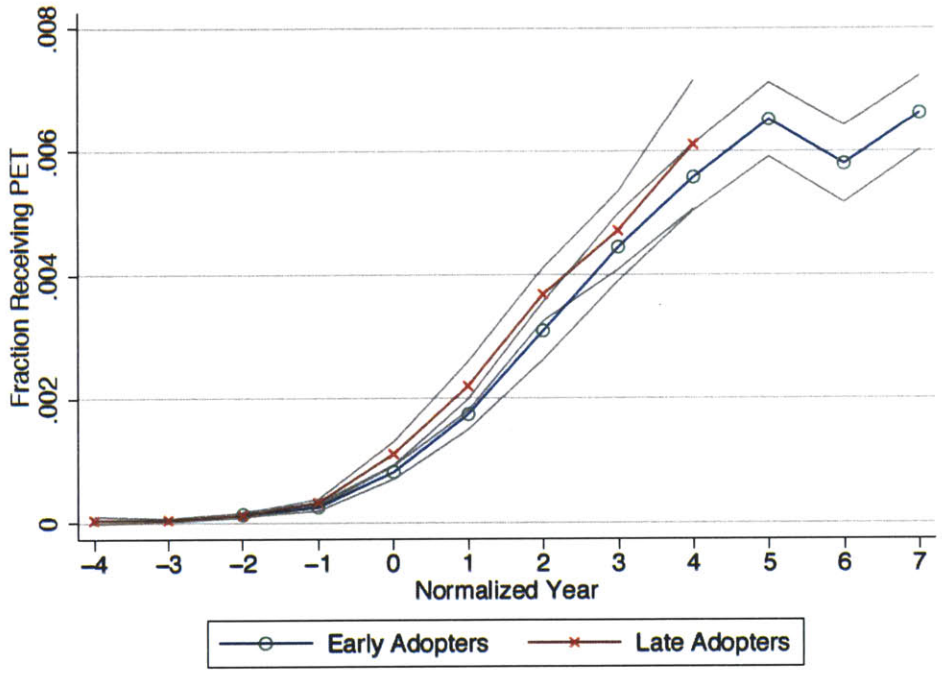


Figure 5: Diffusion of PET Scans in Normalized Time

Notes: This is a plot of state-level diffusion of PET scans as a fraction of the total number of eligible patients. The 95% confidence interval is plotted in grey. Year 0 is the first year in which a state has at least a 0.1% diffusion rate.

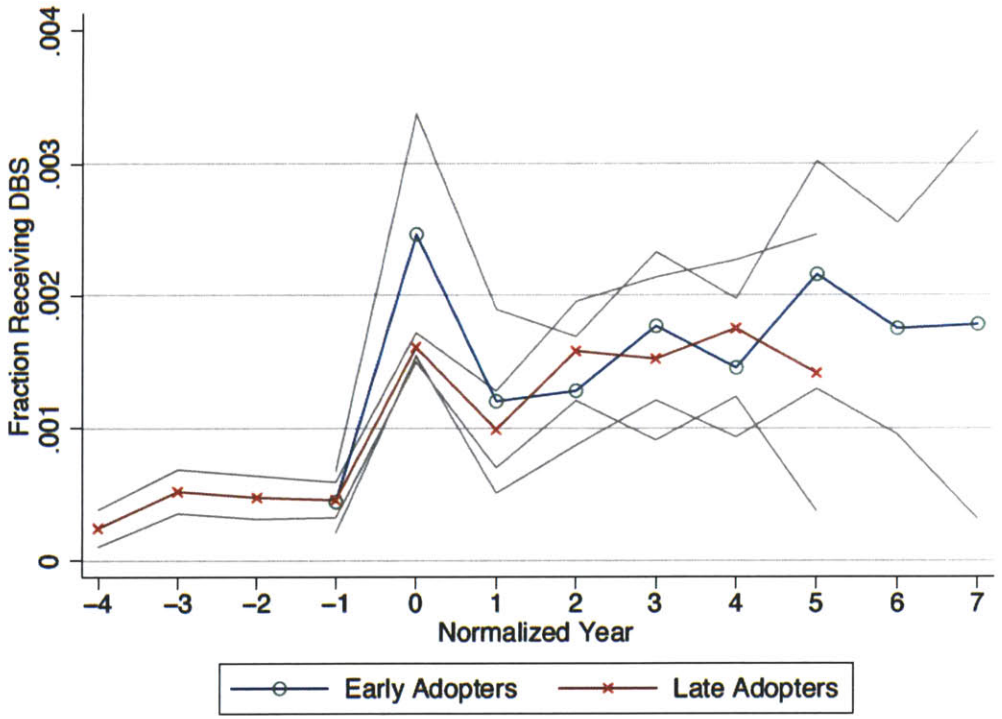
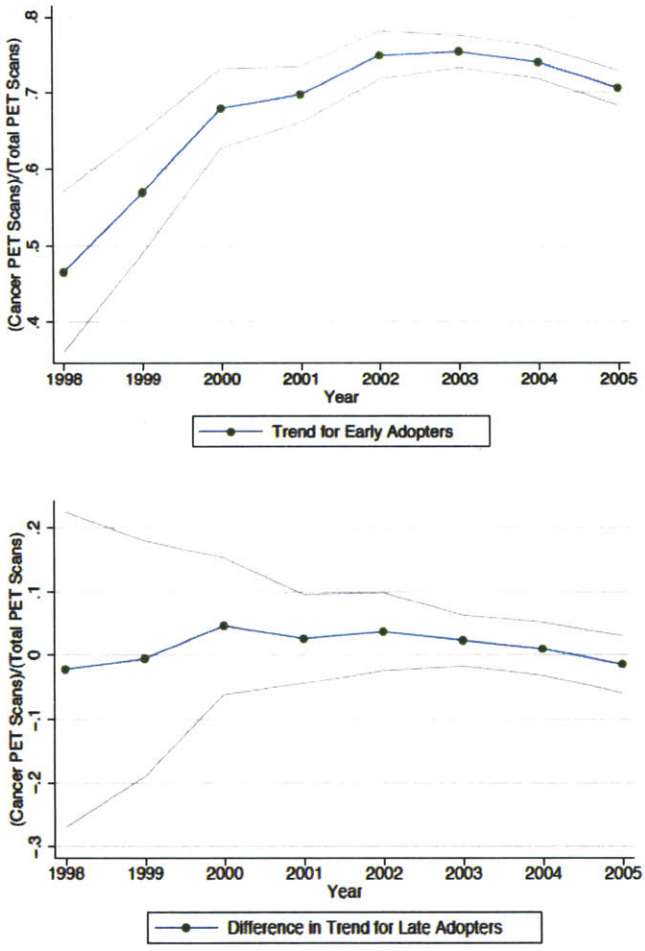


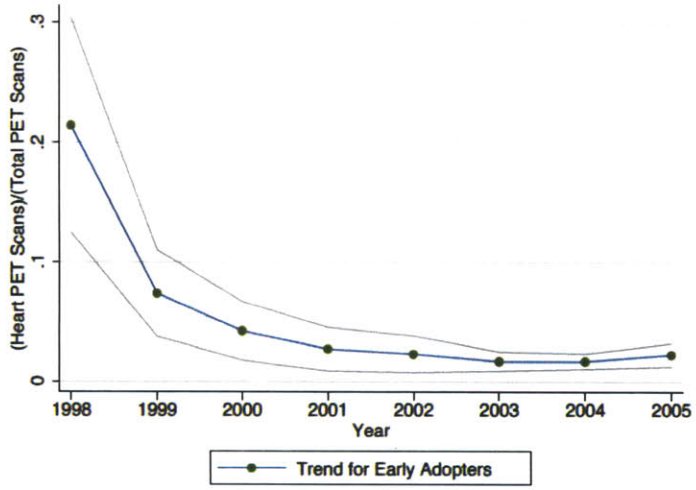
Figure 6: Diffusion of Deep Brain Stimulation in Normalized Time

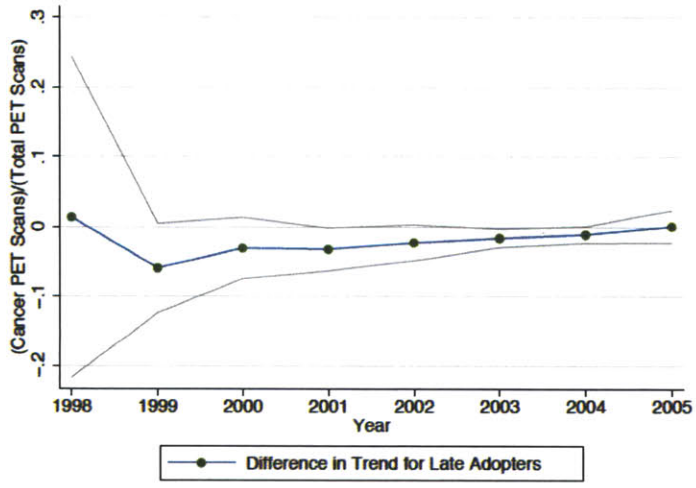
Notes: This is a plot of state-level diffusion of PET scans as a fraction of the total number of eligible patients. The 95% confidence interval is plotted in grey. Year 0 is the first year in which a state has at least a 0.07% diffusion rate.



Figures 7a and 7b: PET Scans for Cancer Indications as a Share of Total Scans.

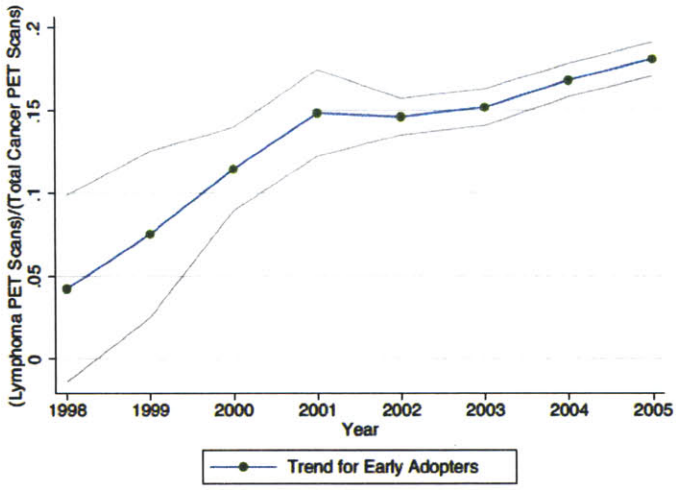
Notes: Figure 7a plots year-by-year average share of cancer scans for early adopters. Figure 7b plots the difference in trend at top experienced by late adopters. 95% confidence interval is in grey

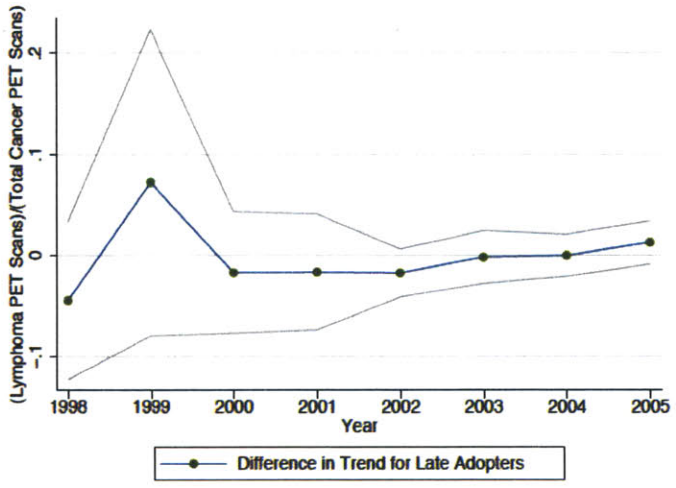




Figures 8a and 8b: PET Scans for Heart Indications as a Share of Total Scans.

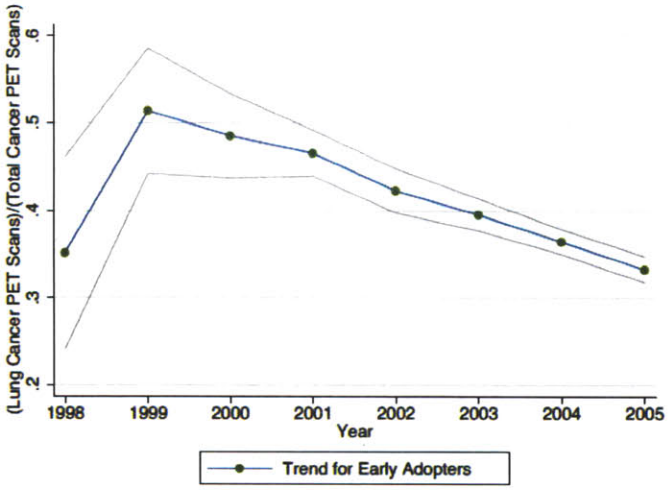
Notes: Figure 8a plots year-by-year average share of heart scans for early adopters. Figure 8b plots the difference in trend at top experienced by late adopters. 95% confidence interval is in grey.

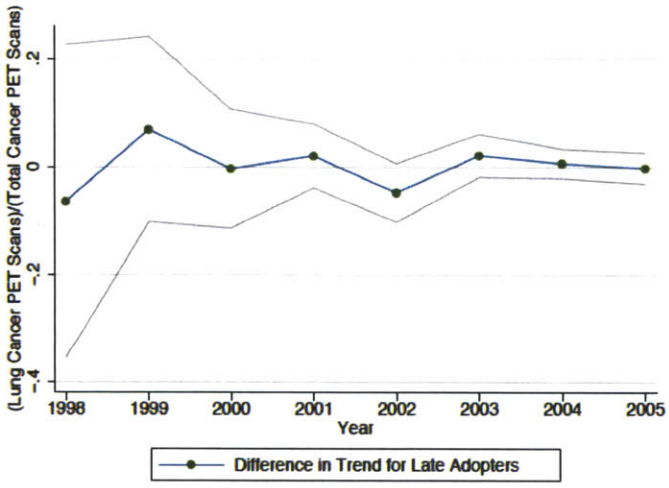




Figures 9a and 9b: PET Scans for Lymphoma as a Share of Total Cancer Scans.

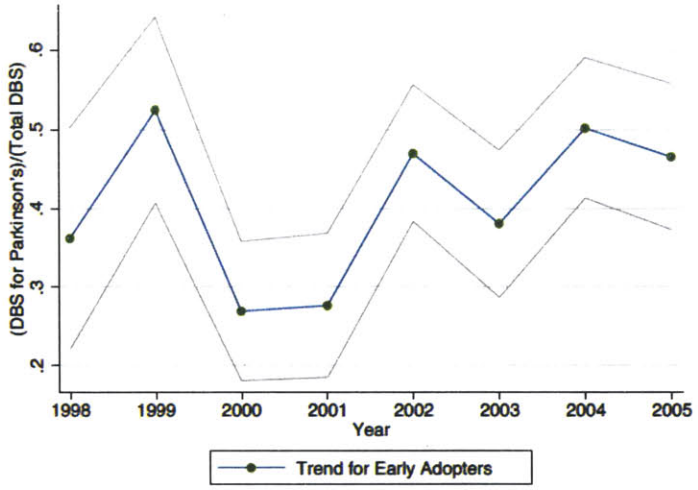
Notes: Figure 9a plots year-by-year average share of lymphoma scans for early adopters. Figure 9b plots the difference in trend at top experienced by late adopters. 95% confidence interval is in grey.

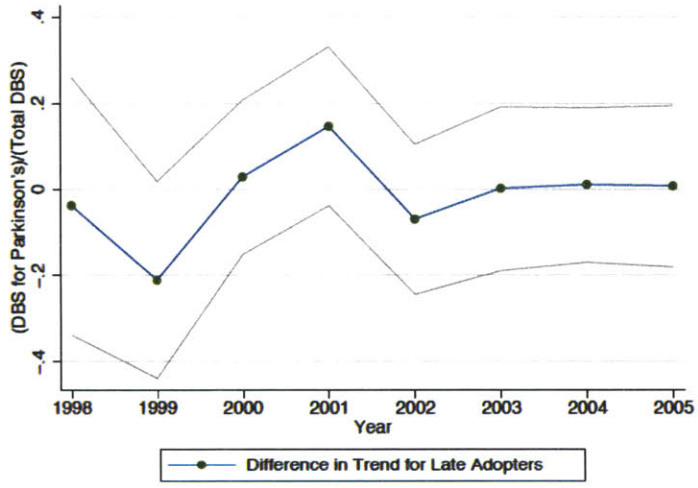




Figures 10a and 10b: PET Scans for Lung Cancer as a Share of Total Cancer Scans.

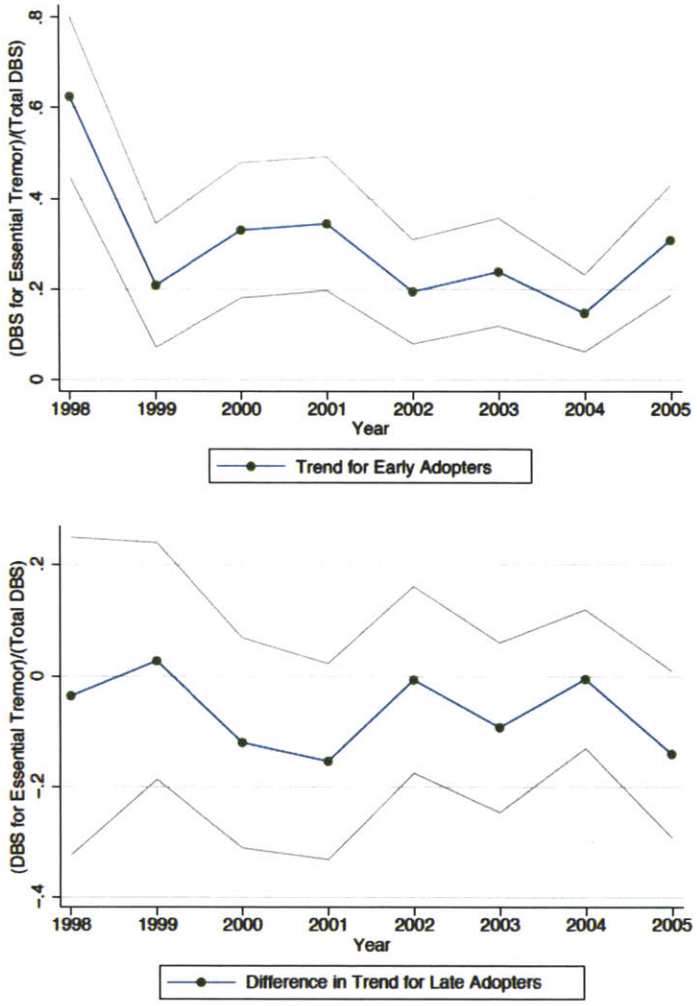
Notes: Figure 10a plots year-by-year average share of lung cancer scans for early adopters. Figure 10b plots the difference in trend at top experienced by late adopters. 95% confidence interval is in grey.





Figures 11a and 11b: DBS Surgeries for Parkinson's as a Share of Total Surgeries

Notes: Figure 11a plots year-by-year average share of Parkinson's surgeries for early adopters. Figure 11b plots the difference in trend at top experienced by late adopters. 95% confidence interval is in grey.



Figures 12a and 12b: DBS for Essential Tremor as a Share of Total Surgeries

Notes: Figure 12a plots year-by-year average share of essential tremor surgeries for early adopters. Figure 12b plots the difference in trend at top experienced by late adopters. 95% confidence interval is in grey.

Table 1: Summary Statistics Contrasting Early and Late Adopting States

	Early Adopters (1)	Late Adopters (2)
A. PET Scans		
No. of states	31	20
No. patients with any eligible diagnosis	147,000	79,600
No. patients with cancer diagnoses	27,800	14,400
1 Year mortality of lung cancer patients	0.048	0.012
Average hospital charges of lung cancer patients	$21,977	$17,682
Average hospital days of lung cancer patients	9.2	8.2
B. Deep Brain Stimulation		
No. of states	27	24
No. patients with any eligible diagnosis	2845	4991
No. patients with Parkinson's diagnoses	1170	1784
1-Year hip fracture rate of Parkinson's patients	0.017	0.020
Average hospital charges of Parkinson's patients	$8,605	$10,074
Average hospital days of Parkinson's patients	3.2	4.2

Notes:

All summary statistics are calculated on an annual basis for the 1998 baseline year.

Sample includes 50 states and the District of Columbia.

Eligible diagnoses include all diagnoses that may be linked to reimbursement for the technology according to Medicare policy.

A state is considered an early adopter of PET scans if they 0.1% of eligible patients received a scan by 2000. A state is considered an early adopter of deep brain stimulation if 0.07% of eligible patients received surgery by 1999.

All data is from 20% sample of Medicare claims.

Table 2: Case Mix of Treated Patients

	Independent Variables			
	Year	Late Adopter	Year*Late Adopter	Multiplicative Model?
Dependent variable				
A. PET Scan Indications				
Cancer scans/Total scans	0.034	0.037	-0.006	No
	(0.006)***	(0.056)	(0.009)	
Cancer scans/Total scans	0.175	0.263	-0.032	Yes
	(0 .030)***	(0.290)	(0.046)	
Heart scans/Total scans	-0.022	-0.047	0.008	No
	(0.047)***	(0.032)	(0.005)	
Heart scans/Total scans	-0.088	-0.178	0.029	Yes
	(0.020)***	(0.129)	(0.021)	
Lymphoma scans/Cancer scans	0.019	0.011	-0.003	No
	(0.003)***	(0.036)	(0.007)	
Lymphoma scans/Cancer scans	0.542	0.247	-0.102	Yes
	(0.087)***	(0.950)	(0.168)	
Lung scans/cancer scans	-0.015	0.026	-0.006	No
	(0.007)**	(0.045)	(0.008)	
Lung scans/cancer scans	0.136	0.416	-0.079	Yes
	(0 .076)*	(0.587)	(0.102)	
B. Deep Brain Stimulation Indications				
Parkinsons DBS/Total DBS	0.009	-0.056	0.012	No
	(0.013)	(0.072)	(0.016)	
Parkinsons DBS/Total DBS	0.024	-0.196	0.019	Yes
	(0.033)	(0.146)	(0.034)	
Essential Tremor DBS/Total DBS	-0.042	-0.116	0.012	Yes
	(0.011)***	(0.066)	(0.012)	
Essential Tremor DBS/Total DBS	-0.392	-0.513	0.149	No
	(0.068)***	(0.470)	(0.076)	

Notes:

The entries report regressions and standard errors (in parenthesis) from 12 separate regressions. The dependent variable is indicated in the leftmost column of each row. Explanatory variables of interest include *Year*, a linear time trend, *Late Adopter*, an indicator for whether the state is a late adopter, and the interaction of the two. The second listed regression for each outcome variable (the Multiplicative Model) includes a control for the indicated diagnoses as a share of the total eligible diagnosed patients, and each included regressor is interacted with the eligible share, as described in equation 1 of the paper.

An observation is a state-year in which at least one application of the technology was observed. Panel A includes 408 observations; Panel B has 333 observations. Standard errors are clustered at the state level.

Asterisks: *10% confidence, **5% confidence, ***1% confidence level.

All data is from 20% sample of Medicare claims.

Table 3: Pre-treatment Health Status of Patients

	Independent Variables		
	Year	Late Adopter	Year*Late Adopter
Dependent variable			
A. Lung cancer patients receiving PET scans			
Pre-period log(hospital charges)	0.011	0.185	-0.029
	(0.012)	(0.100)*	(0.018)
Pre-period days as hospital inpatient	0.099	-0.755	0.124
	(0.076)	(0.474)	(0.094)
Pre-period comorbidity index	0.026	-0.063	0.005
	(0 .0123)**	(0.087)	(0.019)
Standardized Effect	0.035	-0.031	0.002
	(0.017)**	(0.116)	(0.023)
B. Parkinson's patients receiving DBS			
Pre-period log(hospital charges)	0.009	-0.033	-0.015
	(0.033)	(0.131)	(0.030)
Pre-period days as hospital inpatient	0.324	2.825	-0.395
	(0.338)	(2.129)	(0.463)
Pre-period comorbidity index	0.152	3.197	-0.267
	(0.348)	(1.992)	(0.443)
Standardized Effect	0.008	0.122	-0.025
	(0.054)	(0.259)	(0.057)

Notes:

The entries report regressions and standard errors (in parenthesis). The dependent variable is indicated in the leftmost column of each row. There are two observations for each state-year: one for eligible patients and another for patients receiving treatment with the new technology. Explanatory variables of interest include *Year*, a linear time trend, *Late Adopter*, an indicator for whether the state is a late adopter, and the interaction of the two; each of these three variables is equal to zero for all observations representing eligible patients who do not receive treatment with the new technology. Regressions also include controls for year fixed effects that vary according to the state's status as an early or late adopter, and an indicator variable for observations of patients treated with the new technology.

Results reported in the last row of each panel combine estimates from the previous rows to construct a standardized effect.

There are 699 observations for Panel A regressions; 567 observations for Panel B. Standard errors are clustered at the state level.

Asterisks: *10% confidence, **5% confidence, ***1% confidence level.

All data is from 20% sample of Medicare claims.

Table 4: Health Outcomes of Patients Treated with the New Technology

	I. Excluding controls for patient pre-period health: *Independent Variables*			II. Including controls for patient pre-period health: *Independent Variables*		
	Year	Late Adopter	Year*Late Adopter	Year	Late Adopter	Year*Late Adopter
Dependent variable						
A. Lung Cancer Patients Eligible for PET Scans						
Post-period log(hospital charges)	-0.002	0.187	-0.026	0.002	-0.005	0.005
	(0.014)	(.0975)*	(0.021)	(0.017)	(0.064)	(0.013)
Post-period days as hospital inpatient	0.039	0.086	0.020	-0.025	1.077	-0.156
	(0.073)	(0.555)	(0.101)	(0.072)	(0.410)	(0.065)**
Post-period 1 year mortality rate	-0.007	-0.070	0.020	0.017	-0.018	0.007
	(0.009)	(0.044)	(.010)*	(.007)**	(0.057)	(0.014)
Standardized Effect	-0.004	0.045	0.008	0.016	0.098	-0.007
	(0.015)	(0.104)	(0.022)	(0.018)	(0.100)	(0.021)
B. Parkinson's Patients Eligilble for DBS						
Post-period log(hospital charges)	-0.011	-0.227	0.049	-0.008	-0.288	0.048
	(0.020)	(0.154)	(0.031)	(0.023)	(0.202)	(0.034)
Post-period days as hospital inpatient	-0.216	-1.298	0.489	-0.182	-1.847	0.513
	(0.374)	(2.052)	(0.542)	(0.558)	(2.829)	(0.648)
Post-period 1 year hip fracture rate	-0.006	0.006	0.003	0.002	0.032	-0.006
	(0.006)	(0.033)	(0.007)	(0.002)	(0.043)	(0.009)
Standardized Effect	-0.024	-0.193	0.055	-0.006	-0.155	0.030
	(0.028)	(0.173)	0.040	(0.026)	(0.208)	(0.039)

Notes:

The entries report regressions and standard errors (in parenthesis). The dependent variable is indicated in the leftmost column of each row. There are two observations for each state-year: one for eligible patients and another for patients receiving treatment with the new technology.

Explanatory variables of interest include Year, a linear time trend, Late Adopter, an indicator for whether the state is a late adopter, and the interaction of the two; each of these three variables is equal to zero for all observations representing eligible patients who do not receive treatment with the new technology. Regressions also include controls for year fixed effects that vary according to the state's status as an early or late adopter, and an indicator variable for observations of patients treated with the new technology. Panel II regressions include additional controls for patients' health status over the previous year including their inpatient hospital days, hospital spending, and Charlson comorbidity index.

There are two observations for each state-year: one for eligible patients and another for patients receiving treatment with the new technology. There are 671 observations for Panel IA regressions; 605 observations for Panel IB; 584 observations for panel IIA; 474 observations for Panel IIB. Standard errors are clustered at the state level.

Results reported in the last row of each panel combine estimates from the previous rows to construct a standardized effect.

Asterisks: *10 percent confidence, **5% confidence, ***1% confidence

All data is from 20% sample of Medicare claims.

Chapter 3

Managing Medical Technology Policy: A Review of Current Practices and Future Directions [1]

3.1 Introduction

With 32 million Americans projected to gain insurance coverage due to recent health care reform, growing attention is being focused on finding ways to control cost growth

[1]I would like to thank David Autor, Amy Finkelstein, and Michael Greenstone for invaluable guidance throughout this project. I thank Jason Abaluck, Joshua Aronson, Joseph Doyle, Jonathan Gruber, Danielle Li, Amanda Pallais, Heidi Williams, and participants at MIT's labor lunch for their comments. I am also grateful to Mohan Ramanujan and Jean Roth for their assistance in obtaining and managing the data. Arindajit Dube, William Lester, and Michael Reich generously shared data identifying contiguous county pairs. John Wang provided excellent research assistance. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship.

in the health care sector. Medicare and Medicaid currently account for 19% of the federal budget and their fiscal importance is projected to increase over the coming decade. Healthcare spending has been growing at a real rate of about 4.9% per year, and the adoption of new medical technologies has accounted for approximately half of this growth, with demand side factors driving the balance (Congressional Budget Office 2008). As demand for health services continues to rise, the critical challenge lies in managing the utilization of new medical technologies.

Effective technology policy must balance cost control with a recognition that the adoption of new medical technologies has been associated with tremendous health and longevity gains. And yet, the fact that some technologies, including beta blockers, aspirin, amd improved neonatal care, have had high economic rates of returns does not explain the growing puzzle of under- and over-use of many key technologies in the health industry. Some low cost, high value interventions, such as beta blockers, appear persistently underutilized, while some high cost interventions, like cardiac stenting, seem overused despite low marginal returns relative to treatment alternatives.

In this paper, I analyze several policy initiatives that aim to manage the usage of medical technologies and discuss key determinants of technology adoption that may be fruitful targets for future research and policy intervention. In particular, I discuss the process of Medicare coverage for new technologies and the role of state-level certificate of need programs in managing technology diffusion. Together, these constitute a patchwork of programs that are very limited in their ability to manage resource utilization. I then discuss the impact of scientific evidence on technology utilization and coverage policy. Through this discussion, I aim to lay out a framework for understanding medical technology policy in the United States.

Technology adoption in the healthcare sector is distinct from adoption in other

industries in a number of ways, and yet the size and scope of the industry makes it a compelling subject of economic research. Unlike many other technology intensive sectors, e.g. the computer or automative industries, the profit incentives for physicians and hospitals to make efficient adoption decisions are significantly attenuated. In particular, consumers cannot easily observe the quality of the health services and products they are being offered, and relative prices are not always responsive to quality improvements. As a market uniquely plagued by information problems and mis-pricing, the health industry is a critical target for policy aimed to impact technology adoption.

The paper proceeds as follows. Section 2 describes federal and state policies that aim to manage the adoption of new medical technologies. Section 3 describes how scientific evidence is incorporated into medical practice and the formulation of public policy. Section 4 concludes by suggesting directions for future research and what the analyses may be able to say about optimal social policy designed to realize value for healthcare spending.

3.2 Current Technology Policy

The utilization of new medical technologies is currently managed by a patchwork of state, regional, and federal bodies. The result is a diverse but relatively weak set of policy tools aimed at restricting or focusing the adoption of new technologies and managing the market for health services. The programs have some potential for shaping technology usage, but little proven record of performance.

3.2.1 Medicare reimbursement for new technology

As the single largest purchaser of health services in the United States, Medicare is often cited as the industry leader in determining appropriate coverage for new technologies (cf. Van de Water 2010). In principle, Medicare has tremendous power to determine both reimbursement availability and prices for new services. There are two primary channels through which Medicare aims to manage the utilization of new technologies.

First, most new technology policy in Medicare is determined by a set of regional contractors that process claims on behalf of Medicare. Through 2005, this was done by roughly 60 fiscal intermediaries and carriers which processed Medicare Part A and Part B claims, respectively. Physicians were assigned to carriers by geographic regions which roughly correspond to US states, and were required to process Medicare claims through their regional contractor. Hospitals and institutional care providers are allowed to select from a set of potential fiscal intermediaries, but in practice, locally based Blue Cross plans dominate the market, creating regional coverage patterns for Part A claims as well.

Medicare is in the midst of reforming this system of regional contractors and reimbursement, consolidating the role of Part A and Part B contractors into Medicare Administrative Contractors (MACs) which handle both types of claims. Medicare is aiming to reduce the total number of contractors to 10 different MACs nationwide, which are assigned strictly by geographic region.

MACs, fiscal intermediaries, and carriers each have significant latitude to determine "reasonable and necessary" coverage for Medicare recipients. They exercise this judgement primarily through the publication of local coverage decisions which create rules and guidelines designating which medical services are allowable for reim-

bursement for which diagnostic indications. In 2001, soon after the online coverage determination database was first made available, 9000 local policies were posted (Foote 2003). In addition to policy-setting, when a charge is not covered, these contractors are also responsible for the outcome of any appeals process.

The second mechanism for technology management is National Coverage Determinations issued by the Center for Medicare and Medicaid Studies. These rulings are much less frequent, and do not comprehensively guide the set of relevant coverage issues faced by clinicians and contractors. Between 2005-2008, only 13 new policies were published per year, on average (Neumann and Tunis 2010). Even when new policies are published, they are typically broad guidelines that do not specify diagnosis codes or procedure codes, and so local carriers determine the interpretation and implementation of these policies.

Since Medicare is relying heavily on local contractors to determine and enforce coverage rules, it remains a critical question how effective these organizations are at fulfilling this role. I test the impact of local coverage rules on resource utilization by linking Medicare claims data to historical data on regional carrier coverage policies. I investigate the coverage of two technologies, positron emission tomography (PET) and deep brain stimulation, which were the focus of a number of local and national coverage rule changes over the study period from 1998-2005.

Coverage policies were downloaded from the Medicare Coverage Archive Database. For each published policy, the date implemented, covered procedure code, and permissible diagnosis codes are documented. Thus, I constructed a novel data set of the evolution of coverage across regions, over time, for these two technologies.

Data on utilization comes from a 20% sample of Medicare Part B claims. These include claims rendered for physician services delivered in both an inpatient and outpatient context. Using this data, I construct a national panel data set on deep

brain stimulation and PET scan usage, which is linked to the reimbursement regime of the corresponding local Medicare carrier.

PET scanning is a nuclear imaging technique that produces a three dimensional image of cellular metabolic activity. The technology is now widely used in clinical oncology, and in more limited use to diagnose certain neurologic diseases and to map heart function. PET scans were the subject of dozens of local coverage policies over the study period, as well as a series of National Coverage Determinations. First approved for applications to heart disease, PET became an increasingly popular imaging tool for a variety of cancer diagnoses. In successive waves, local contractors and eventually the national Center for Medicare & Medicaid Studies, expanded the covered indications.

I analyze the impact of these PET coverage policies on three indications that had a particularly high variance in the timing of coverage : esophageal cancer, nervous system disorders, and breast cancer. For each of these indications, between 7 and 11 early-coverage carriers had provided reimbursement by 1999, and between 55 and 59 remaining carriers did not issue a ruling until after a National Coverage Determination was made in 2001 (for esophageal and nervous system indications) or 2002 (for breast cancer).

The second technology under analysis, deep brain stimulation, is a surgical treatment for neurological movement disorders. The procedure involves the implantation of a device that sends electrical impulses to targeted areas of the brain, suppressing involuntary movement. Deep brain stimulation was also the subject of many coverage policies over the study period. 19 carriers provided coverage within two years of the 1997 FDA approval; an additional 45 carriers did not issue coverage until after a National Coverage Determination was issued in 2003.

I test the impact of coverage policy on utilization in two ways. First, I plot diffu-

sion curves and corresponding 95% confidence intervals for early- and late-covering carriers, allowing a visual inspection of how diffusion evolves after a significant coverage expansion. Figures 1 through 4 display these results.

In Figure 1, it appears that the application of PET scans to esophageal cancer diffused more quickly in early coverage regions, but the difference between the two groups becomes most striking *after* 2001, the year in which all carriers expanded coverage in compliance with a National Coverage Determination.

Despite the coverage changes over this period, utilization of PET scans for disorders of the nervous system remained rare amongst both early- and late-coverage regions. As illustrated in Figure 2, the two diffusion curves are not statistically distinguishable from each other.

In Figure 3, there is very little difference between early- and late-coverage areas in the diffusion of PET scans for breast cancer, and no indication of a change in the diffusion patterns in 2002, when late adopters gained reimbursement permission.

Lastly, in Figure 4, diffusion of deep brain stimulation is similar across early and late coverage areas, and there are no differential changes in diffusion after the 2003 coverage of late adopters.

One possible explanation for the observed lack of correlation between coverage policy and the carrier-level diffusion curves is that heterogeneity in the patient population or physician characteristics across regions may drive the diffusion patterns, masking the actual impact of coverage policy. For this reason, I also implement a strategy modeled after Dube, Lester, and Reich (2010) that compares early and late adopters across contiguous counties, reducing the population heterogeneity of the comparison groups.

For the border county comparisons, I limit the analysis to contiguous county pairs that rest along state borders assigned to different contractors with disparate

coverage policies. Regressions take the following form:

$$\text{ShareTreated}_{ipt} = \alpha_i + \beta_{pt} + \gamma_1 \text{CoveragePolicy}_{it} + \gamma_2 \text{CoveragePolicy}_{it}\text{Year}_{it} + \epsilon_{ipt} \tag{3.1}$$

Where $\text{ShareTreated}_{ipt}$ is the fraction of potentially eligible patients receiving the new treatment in county i that is part of pair p at time t. The regression includes county fixed effects α_i and pair-specific year fixed effects β_{pt}. The independent variables of interest are $\text{CoveragePolicy}_{it}$ and $\text{CoveragePolicy}_{it}\text{Year}_{it}$. These variables indicate whether the county has an applicable coverage policy in place and allow the time trend for the county pair to diverge for the adopting county after the implementation of a policy. Because each county may be paired with more than one neighbor, county-year observations may appear multiple times in the data set, and standard errors are clustered at the county level.

Results from these regressions are reported in Table 1 Panel A. There is no evidence for any of the technology applications of significant increases in utilization after a coverage expansion. In fact, the point estimates are negative in most specifications, and for PET scans for nervous system disorders the decline in utilization is statistically significant at the 5% level. This result corroborates the more aggregate graphical evidence that the implementation of new coverage policies was not associated with changes in the utilization of a new medical technology.

The findings suggests that Medicare coverage rules are not effective at constraining the utilization of new technologies. This finding is consistent with evidence from Foote et al. (2008) which tested the impact of Medicare coverage policies for eight technologies using contractor-level comparisons. Foote and Town (2007) argue that Medicare contractors lack both resources and incentives to enforce the stated coverage policies. In several of the cases analyzed by Foote et al., contractors would be

required to audit the patient's chart to verify if the claim was compliant with the stated guidelines. However, even in the cases I analyze above, where the listing of a covered diagnosis code on the claim should be a necessary condition for providing coverage, there is no evidence that carriers were rejecting claims without the requisite coverage. Perhaps more critically, contractors are awarded contracts on the basis of low overhead on a per-claim basis, not for appropriate enforcement of coverage policy (Foote and Town 2007).

There are policy tradeoffs between the costs of auditing and the savings associated with better enforcement, but even simple computer algorithms could identify submitted claims for broad diagnostic indications that are not covered by current policy. Whether due to inattention or unwillingness, Medicare does not seem to enforce effectively evidence-based policy constraints that might interfere with provider judgement about appropriate medical care. This is one arena in which clearer enforcement of existing policy could have significant potential to shape the utilization of new medical technologies. The potential impact of policy enforcement may even have increased over the past decade, as Medicare's National Coverage Determinations have become more specific in citing a lack of appropriate scientific evidence as motivation to restrict or deny coverage for a new technology (Neumann and Tunis 2010). It is possible these practice guidelines have a small and diffuse effect on medical practices, but they seem to have little bite in quickly or significantly impacting resource utilization.

3.2.2 Certificate of need programs

A second existing policy mechanism for controlling the adoption of new medical technologies is Certificate of Need (CON) programs. These programs are intended

to reduce the supply of capital-intensive health services by requiring state approval for large capital projects such as building expansions or new high-tech devices. Many CON programs originated in response to the federal Health Planning Resources Development Act of 1974, which required each state health planning agency to have a structure for approving large capital projects, and provided federal funding to CON programs. The federal mandate and funding was repealed in 1987, but as of 2008, 36 states retained some form a CON program (National Conference of State Legislatures 2010).

CON programs aim to control utilization of medical services by restricting market entry, rather than restricting reimbursement payment for services. Notably, the most common political argument in favor of CON regulation is that "excess capacity (in the form of facility overbuilding) directly results in health care price inflation" (National Conference of State Legislatures 2010), contrary to basic economic models that would predict falling prices as supply competition increases.

The most frequent medical technologies to full under CON jurisdiction are cardiac catheterization (26 states), CT scanners (15), gamma knives (19), lithotripsy (21), MRI scanners (21), and PET scanners (23). As can be seen on the map in Figure 5, CON programs are concentrated along the Eastern seaboard and the Midwestet.

There is little recent evidence on how the presence of a CON program affects technology investment and usage. Conover and Sloan (1998) find that CON programs are not associated with major differences in the concentration of technology-intensive open heart surgery or organ transplant units. To further investigate the relationship between CON programs and the diffusion of capital-intensive new medical services, I examine whether the diffusion of PET scans is slower in states covered by germane CON programs. The analysis utilizes the 20% sample of Medicare claims from 1998-2005 merged with data on CON program coverage from the National Conference of

State Legislatures.

In Figure 6, the fraction of eligible patients receiving a PET scan is plotted for states with and without CON programs. There is no visible evidence of slower diffusion amongst states with applicable CON policies, suggesting another instance of a technology utilization management policy that fails to have any measurable impact on clinical practice.

In addition to the graphical evidence described above, I estimate regressions of the following form:

$$\text{ShareTreated}_{st} = \alpha_t + \beta_1 \text{CONState}_s + \beta_2 \text{CONState}_s \text{Year}_t + \epsilon_{st} \qquad (3.2)$$

The dependent variable is the fraction of eligible patients treated with a PET scan in state s in year t. Included are year fixed effects α_t, and the independent variables of interest are an indicator for whether the state has a CON program and an interaction term between the presence of a CON program and the time trend.

The findings reported in Table 1 Panel B show no evidence of slower diffusion amongst CON program participants. The coefficient estimates on the CON policy indicator and the interaction between CON policy and time are both positive, albeit very small in magnitude, and not statistically distinguishable from zero. State CON programs that covered the adoption of PET scanners appear to have had no effect on PET scanning rates over this period.

One limitation of this analysis is that since indications for the usage of PET scanners were continually expanding over this period, it is possible that the CON programs assessed that their states had no excess supply of scans over this period. If this is the case, CON may have a greater impact on more mature technologies, although if the CON programs do not slow the entry of new providers at some point

during the diffusion process, they would have to rely on firm exit to enforce future reductions in the target number of providers in an area.

A second limitation is that the above regression does not exploit panel variation in the presence of CON programs, and cannot separately identify differences in physician preferences and patient characteristics across states from the impact of the CON program. The regression above provides suggestive evidence that CON does not lead to substantially lower levels of technology usage, but it cannot rule out the possibility that CON states would have had otherwise higher levels of PET scan utilization than their non-CON peers.

The evidence on the impact of CON programs echoes the findings that stated Medicare coverage policies have little effect on the utilization of new technologies. Despite the energy and attention given to these regulations, it seems that the current set of policies do not have a substantial impact on the application of new technologies in practice. Given the growing costs and uneven value realized from technology adoption, further attention should be given to either improving enforcement of current policies, or improving the structures that support learning about the value and applications of a new technology.

3.3 From Scientific Evidence to Clinical Practice

3.3.1 Impact of new research on physician behavior

The unique challenge to constructing efficient policy for managing the usage of new technologies is that the scientific understanding of the technology's efficacy and role relative to alternative treatment options is continuously evolving. Existing structures for discovering and disseminating new information about a technology's value and

applicability are limited in their ability to manage technology usage. There are two critical issues: first, how quickly and effectively physicians can learn and incorporate new information and services into their clinical practice; second, how policy can facilitate or improve the efficiency of that process.

On the question of how physicians incorporate new scientific information, there is limited existing evidence. Azoulay (2002) finds that physician prescribing patterns do respond to new scientific evidence, and that marketing efforts to disseminate favorable information intensify these effects, increasing the total responsiveness. This suggests a complementarity between scientific discovery and marketing that may be particularly valuable for innovations with close competitors in the market, and a firm likely to profit directly from disseminating the information.

Physicians seem to react much less than would be socially optimal to the accumulation of new evidence about the value of off-patent medications. For example, several studies have documented the persistent under-prescription of antithrombotic therapy, such as aspirin, after a stroke or heart attack despite well established evidence-based criteria and low costs of the intervention (Gage et al. 2000, Califf et al. 2002). If the evidence from Azoulay (2002) is applicable to other classes of drugs and types of interventions, it suggests there could be social gains from public investment in marketing messages targeted to physicians regarding appropriate application of highly effective but under-used treatments. For interventions where the profit incentive to disseminate scientific information is lacking, subsidizing information acquisition by clinicians and patients may be a worthwhile investment. Further investigation is warranted into the relative costs and efficacy of price mechanisms that provide direct incentives to align behavior with evidence-based recommendations and information interventions that seek to educate physicians about appropriate treatment interventions.

The overuse of medical interventions which have been shown to have low clinical value also requires research and policy attention. The canonical example is the placement of cardiac stents, and drug-eluting stents in particular (cf. Mitka 2006). Although some patients benefit notably from this intervention, the treatment seems to have low returns on the margin for some patients who may otherwise have received medical management, rather than a surgical intervention. The proliferation of treatment by stenting may be driven by a number of factors: financial incentives of physicians to bill for the additional procedure, uncertainty about the medical returns to intervention, or consumer demand for the treatment.

Although physician incentives in a fee for service setting is a key ingredient, it is likely not be the sole driver of overuse. Even in a managed care setting, a Kaiser Permanente executive argued that Kaiser must over-provide stenting services, beyond what would be clinically optimal, to avoid losing patients to other practices. Dr. Carl Weisberger argued "around 70% of people getting an angioplasty don't need a stent" and yet purchasing stents for the practice is "worth it because the community perception demands you use the item" (Pope 2006). This observation suggests a role for market competition and consumer demand in the over-provision of medical services. Are managed care organizations facing fewer competitive pressures better at reducing the provision of unproductive services? Are there ways to directly inform consumers that will reduce their demand for services that are not clinically indicated?

Marketing to both patients and physicians has shown to be effective at increasing take-up of certain indicated drugs and interventions. Surveyed physicians report that direct to consumer advertising helps identify previously undiagnosed conditions and facilitates awareness and discussions of treatment options (Berndt 2005). If targeted marketing can mitigate the problem of *underuse* of appropriate interventions, can marketing also be used to target *overuse*? In light of the apparent difficulty or lack

of will to enforce policies designed to prevent overuse, as discussed above, further investigation is warranted into whether informational nudges, rather than direct restrictions, can be used to reduce demand for over-used services.

3.3.2 Impact of scientific evidence on coverage rules and policy setting

The formulation of public policy must also grapple with the changing landscape of scientific information. Setting aside the enforcement issues associated with Medicare coverage rules, the Center for Medicare and Medicaid Studies has spent an increasing amount of resources analyzing, documenting, and discussing how coverage policy responds to scientific evidence. Neumann and Tunis (2010) document a trend in National Coverage Determinations, which are increasingly ruling to limit coverage of a new technology in cases where the existing evidence has a "lack of relevant outcomes" or "lack of applicability to the Medicare population." The authors suggest the burden of proof is shifting for Medicare coverage to a presumption of non-coverage unless scientific evidence of improved health outcomes has been robustly documented.

The move towards more restrictive coverage of new technologies introduces trade-offs between physician autonomy, which may rely on a mix of scientific evidence and experiential learning to determine a clinical course, and external rulings which rest exclusively on scientific learning. It is possible that this could have a distributional impact on the quality of care delivered that depends both on the quality of the physician and the disease severity or rarity of the patient's condition. Rigid rules could cause the performance of the best physicians to deteriorate or harm the care of the most unusual patients, while improving outcomes near the means of the distribution.

Some case study evidence on the enforcement of strict practice guidelines is avail-

able from Intermountain Health Care, and shows generally positive impact on health outcomes (Dean et al. 2001). If more restrictive reimbursement rules become enforced in other contexts, further study on the tradeoffs between autonomy and rule-based practices is warranted.

Relatedly, there has been a small but growing number of technologies that become approved for Medicare coverage with evidence development, in cases where a novel treatment seems promising enough to warrant coverage but there is a lack of reliable data on the clinical performance and outcomes. The first case of coverage with evidence development was lung volume reduction therapy. This treatment for emphysema was first developed in the 1950s, but it exploded in popularity the mid-1990s despite a lack of robust scientific evidence. The intervention required no new capital investments or training, and made use of existing thoracic surgery techniques to remove diseased lung tissue. Physicians submitted claims for the surgery using existing billing codes, and within 18 months of the inciting clinical case study, 1200 surgeries had been performed on Medicare patients (Ramsey and Sullivan 2005).

Observing the growing volume of surgeries and lack of clinical evidence, the Center for Medicare and Medicaid Studies developed a policy whereby it would reimburse the surgery only for patients enrolled in a randomized controlled trial. The results of the trial were nuanced and showed modest benefit of the surgery for only some subgroups of potentially eligible patients. Moreover, the intervention had poor cost-effectiveness. Despite these findings, Medicare approved coverage for a relatively broad set of indications in 2004, although only at sites that had participated in the clinical trial or were designated lung transplant centers. An interesting direction for future research would be contrasting the volume of lung volume reduction therapy at the clinical trial hospitals with other transplant centers to see how a physician's own experience with a technology may mediate his reaction to new scientific information.

After this experiment with lung volume reduction therapy, Medicare developed a new program for coverage with evidence development in 2005, and has since provided coverage for several new technologies under the program (Tunis and Pearson 2006). This represents a new effort to develop reimbursement rules that are responsive scientific measurement of a technology's value. The success of this effort may depend on the willingness to restrict coverage if the results of the study are not favorable—the road not taken in providing coverage for lung volume reduction therapy—as well as Medicare's ability to enforce its own coverage determinations.

3.4 Conclusion

This paper highlights many of the challenges inherent in developing effective public policy for the management of new medical technologies. Existing policies have little demonstrated capacity to shape medical practice. Medicare coverage policies for new technology have no measurable impact on utilization rates, despite the significant resources devoted to drafting coverage policies. Contractors that process claims for Medicare have little incentive to devote resources to enforcing coverage rules, since contracts are awarded primarily on the basis of the contractor's overhead costs per claim processed. If Medicare would like to take a more active role in managing the utilization of medical resources, it may need to start with an investigation of how to improve compliance with its current recommendations.

Many states make use of certificate of need programs to reduce the supply of capital-intensive medical services, with the aim of reducing health care costs. These programs have little proven track record for reducing resource utilization, and I find no cross-sectional correlation between the presence of a CON program that manages PET adoption and the rate of PET scan utilization in that state. In addition, there is

relatively scant evidence that oversupply of medical services raises costs or generates inefficiencies. On the one hand, Gruber and Owings (1994) find that in areas where obstetricians experience a drop in demand, they increase the number of cesarean sections performed to compensate with higher revenue per patient for the drop in total patient volume. On the other hand, competition in the market for medical services may improve the quality of services offered by providing incentives for better management practices (Bloom et al. 2010). Without a richer understanding of the tradeoffs associated with regulating the supply of medical services, it is unclear what the optimal role of a CON program would be. Moreover, a CON program can only influence supply of capital-intensive technologies; while this is a critical component of new medical technology, CON programs alone are clearly cannot provide comprehensive solutions to resource management.

Together, the evidence on Medicare coverage and CON programs suggest that there is significant scope for expanding the role of regulation in managing the adoption of new medical technologies. However, if these policy programs were to take a more active role in limiting access to new technology, the burden becomes much greater to select the optimal levels of coverage and ensure that there is sufficient evidence available to guide the formulation of coverage policy. Medicare has taken steps in this direction, developing a new program to facilitate coverage with evidence development that provides reimbursement for particular services only if physicians participate in a clinical study. It will take more comprehensive reform then the implementation of coverage with evidence development policies though, if Medicare would like to take a more active role in resource utilization. It would need to not only use the generated evidence to limit reimbursement for technologies that prove to be low value, essentially taking away coverage for indications that were previously paid for by the program, but also to reform enforcement efforts to ensure that the

policies it sets have an impact on clinical practice.

Policymakers have a second set of tools available, beyond price incentives, coverage limitations, and supply regulation, for managing the utilization of new medical technologies. Specifically, both physicians and patients have been shown to be responsive to marketing information about pharmaceutical treatments, and for physicians in particular, marketing complements the publication of scientific information. This suggests that subsidizing information acquisition through marketing efforts may provide another avenue for encouraging desirable clinical practices. The existing evidence largely consists of encouraging the use of particular pharmaceuticals; future research could investigate whether it is possible for informational nudges to reduce utilization of low-value services. In addition, an evaluation that contrasts the benefit and cost of information provision with the benefit and costs of setting and enforcing coverage policy would provide a useful point of comparison.

As more Americans gain health insurance and the costs of care continue to rise, managing the utilization of new medical technologies is of growing importance. There is scant evidence that the current structures for managing the adoption of new technologies, including Medicare coverage policy and state certificate of need programs, are efficient or even effective at enforcing their stated policy goals. Setting policy to manage technology adoption requires a fine balance between facilitating valuable experimentation and learning about new interventions, encouraging responsiveness to scientific evidence, and maintaining physician autonomy when appropriate. Continued research into how physicians respond to reimbursement rules, price incentives, and new information will be crucial to guiding policy reform that can achieve that balance.

3.5 References

Azoulay, Pierre. 2002. "Do Pharmaceutical Sales Respond to Scientific Evidence?" *Journal of Economics and Management Strategy*, 11(4): 551–594.

Berndt, Ernst R. 2005. "To Inform or Persuade? Direct-to-Consumer Advertising of Prescription Drugs." *The New England Journal of Medicine*, 352(4): 325–329.

Bloom, Nicholas, Carol Propper, Stephan Seiler, and John Van Reenen. 2010. "The Impact of Competition on Management Quality: Evidence from Public Hospitals." NBER Working Paper No. 16032.

Califf, Robert M, Elizabeth R DeLong, Truls Ostbye, Lawrence H Muhlbaier, Anita Chen, Nancy Allen LaPointe, Bradley G Hammill, Charles B McCants, and Judith M Kramer. 202. "Underuse of aspirin in a referral population with documented coronary artery disease." *American Journal of Cardiology*, 89(6): 653–661.

Cauchi, Richard. 2009. "Certificate of Need: State Health Laws and Programs." National Conference of State Legislatures.

Congressional Budget Office. January 2008. "Technological Change and the Growth of Health Care Spending."

Conover, Christopher J., and Frank A. Sloan. 1998. "Does Removing Certificate-of-Need Regulations Lead to a Surge in Health Care Spending?" *Journal of Health Politics, Policy and Law*, 23(3): 455–481.

Dean, NC, MP Silver, KA Bateman, B James, CJ Hadlcok, and D Hale. 2001. "Decreased mortality after implementation of a treatment guideline for community-acquired pneumonia." *American Journal of Medicine*, 110(6): 451–457.

de Water, Paul N. Van. 2008. "Medicare Change Can Complement Health Reform." Center on Budget and Policy Priorities.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum Wage Estimates Across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics*, 92(4): 945–964.

Foote, Susan Bartlett. 2003. "Focus On Locus: Evolution Of Medicare's Local Coverage Policy." *Health Affairs*, 22(4): 137–146.

Foote, Susan Bartlett, and Robert J. Town. 2007. "Implementing Evidence-Based Medicine Through Medicare Coverage Decisions." *Health Affairs*, 26(6): 1634–1642.

Foote, Susan Bartlett, Beth A. Virnig, Robert J. Town, and Lacey Hartman. 2008. "The Impact of Medicare Coverage Policies on Health Care Utilization." *Health Services Research*, 43(4): 1285–1301.

Gage, Brian F., Michael Boechler, Amy L. Doggette, Gary Fortune, Greg C. Flaker, Michael W. Rich, and Martha J. Radford. 2000. "Adverse Outcomes and Predictors of Underuse of Antithrombotic Therapy in Medicare Beneficiaries With Chronic Atrial Fibrillation." *Stroke*, 31: 822–827.

Gruber, Jonathan, and Maria Owings. 1994. "Physician Financial Incentives and Cesarean Section Delivery." NBER Working Paper No. w4933.

Mitka, Mike. 2006. "Critics Say Drug-Eluting Stents Overused." *Journal of the American Medical Association*, 296(17): 2077.

Neumann, Peter J., and Sean R. Tunis. 2010. "Medicare and Medical Technology — The Growing Demand for Relevant Outcomes." *New England Journal of Medicine*, 362(5): 377–379.

Pope, Tom. 2006. "Hard to Curb Overuse of Drug-Eluting Stents." *Managed Care.*

Ramsey, Scott D., and Sean D. Sullivan. 2005. "Evidence, Economics, And Emphysema: Medicare's Long Journey With Lung Volume Reduction Surgery." *Health Affairs*, 24(1): 55–66.

Tunis, Sean R., and Steven D. Pearson. 2006. "Coverage Options For Promising Technologies: Medicare's 'Coverage With Evidence Development'." *Health Affairs*, 25(5): 1218–1230.

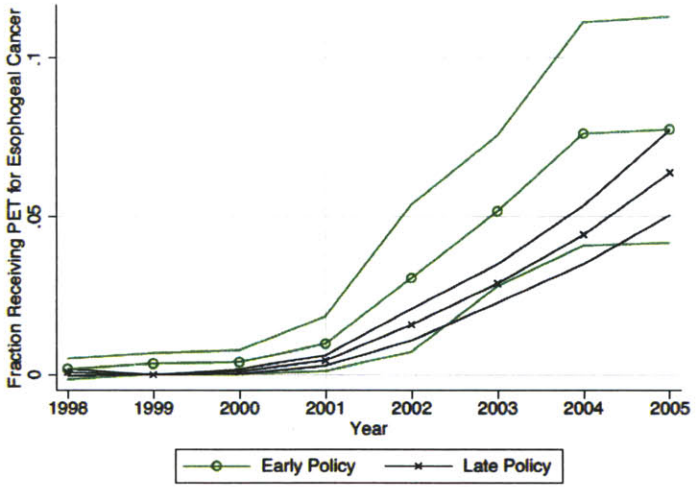


Figure 1: Diffusion of PET Scans for Esophageal Cancer

Notes: This is a plot of carrier-level diffusion of PET scans as a fraction of the total number of eligible patients. Early Policy carriers have put a coverage policy in place by 1999; Late Policy carriers do not have a coverage policy until 2001. The 95% confidence interval is plotted for each curve in the corresponding color.

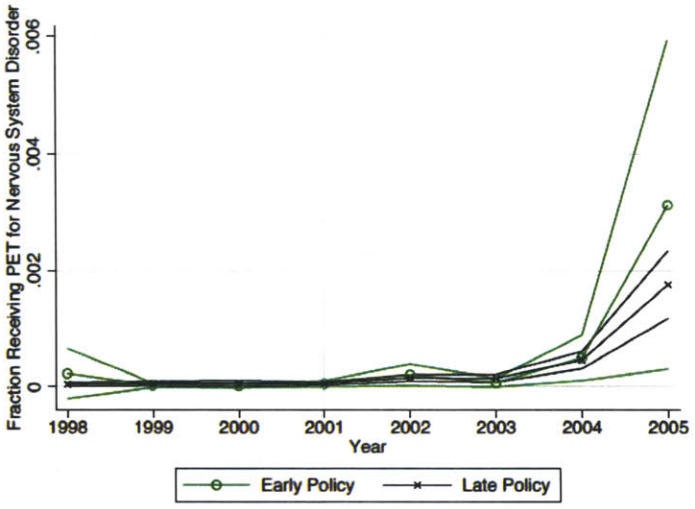


Figure 2: Diffusion of PET Scans for Nervous System Disorders

Notes: This is a plot of carrier-level diffusion of PET scans as a fraction of the total number of eligible patients. Early Policy carriers have put a coverage policy in place by 1999; Late Policy carriers do not have a coverage policy until 2001. The 95% confidence interval is plotted for each curve in the corresponding color.

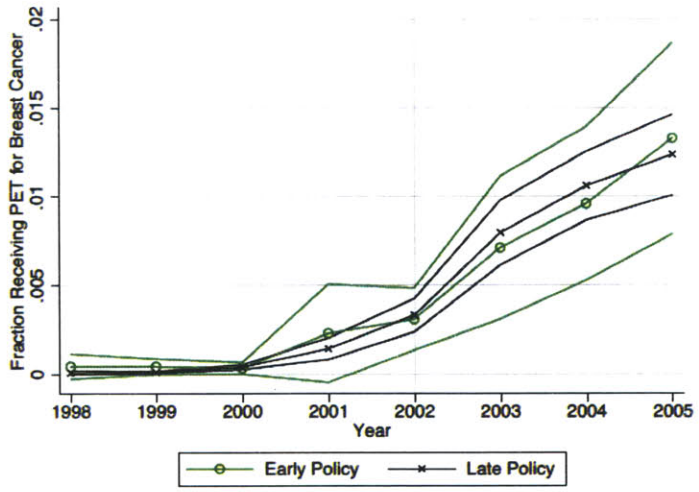


Figure 3: Diffusion of PET Scans for Breast Cancer

Notes: This is a plot of carrier-level diffusion of PET scans as a fraction of the total number of eligible patients. Early Policy carriers have put a coverage policy in place by 1999; Late Policy carriers do not have a coverage policy until 2002. The 95% confidence interval is plotted for each curve in the corresponding color.

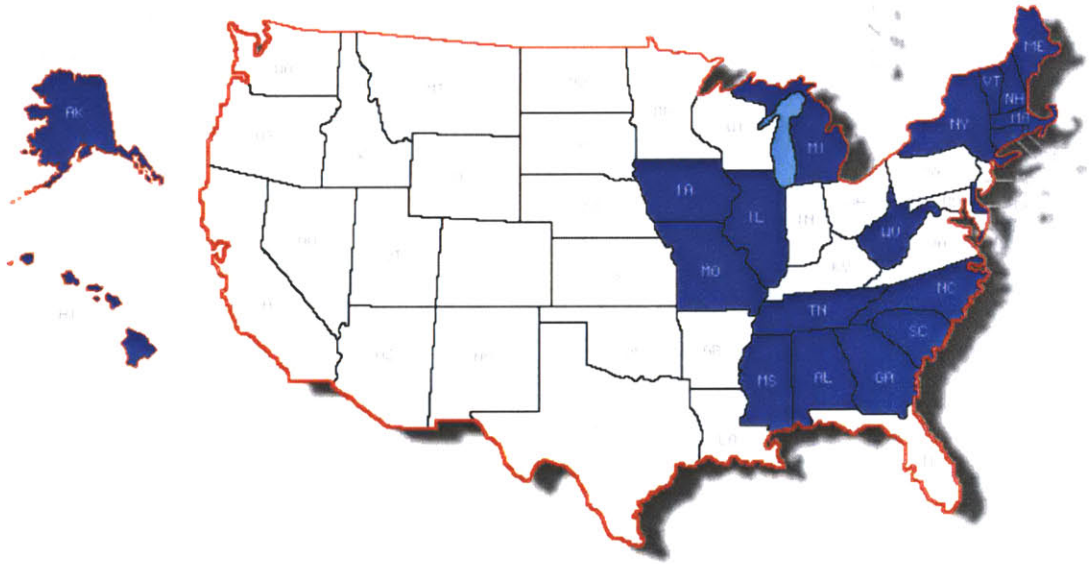


Figure 5: Map of Certificate of Need Programs Covering PET Scans

Note: States with CON programs are highlighted in blue.

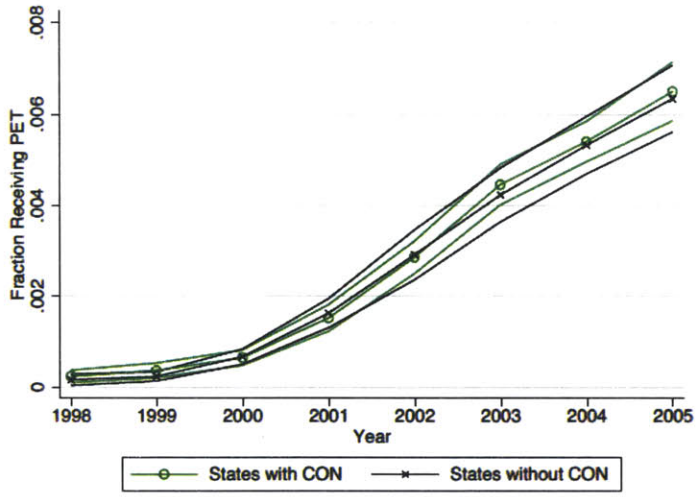


Figure 6: Diffusion of PET Scanners

Notes: This is a plot of state-level diffusion PET scanners as a fraction of the total number of eligible patients. Diffusion curves are plotted separately for states with and without Certificate of Need programs. The 95% confidence interval is plotted for each curve in the corresponding color.

Table 1: Impact of Technology Policy on Utilization of Health Services

Dependent Variables	*Independent Variables*		*No. of Obs.*
A. Impact of Medicare Coverage Policies			
	Coverage Policy	Year*Coverage Policy	
PET Scans for Breast Cancer	0.00069	-0.00018	8518
	(0.00047)	(0.00028)	
PET Scans for Central Nervous System	-0.00033	0.00011	8563
	(.00016)**	(0.00007)	
PET Scans for Esophogeal Cancer	-0.00801	-0.00159	6171
	(0.00750)	(.00083)*	
Deep Brain Stimulation	-0.00052	-0.00016	8081
	(0.00073)	(0.00017)	
B. Impact of State Certificate of Need Programs			
	CON Policy	Year*CON Policy	408
Total PET Scans	0.00001	0.00001	
	(0.00014)	(0.00007)	

Notes:

The entries report regressions and standard errors (in parenthesis). The dependent variable is indicated in the leftmost column of each row. The dependent variables of interest are an indicator for whether the relevant policy is in place, and an interaction of the indicator with a linear time trend.

Panel A: an observation is a county-year, and the sample includes contiguous counties that have different carrier assignements and thus face different coverage policies. Panel B: an observation is a state-year.

Panel A: Regressions include controls for county pair specific year fixed effects. Standard errors are clustered at the county level. Panel B: Standard errors are clustered at the state level.

Asterisks: *10% confidence, **5% confidence, ***1% confidence level.

All data is from 20% sample of Medicare claims.