Essays on the Impact of Health Information Technology on Patient Outcomes

A Dissertation presented

by

Ryan Michael McKenna

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Economics

Stony Brook University

May 2016

Stony Brook University

The Graduate School

Ryan Michael McKenna

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Dr. Debra Sabatini Dwyer
Associate Professor, Department of Technology and Society

Dr. Mark Montgomery
Associate Professor, Department of Economics

Dr. John Rizzo
Associate Professor, Department of Economics

Dr. Pamela Noack
Clinical Assistant Professor, Department of Technology and Society

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Essays on the Impact of Health Information Technology on Patient Outcomes

by

Ryan Michael McKenna

Doctor of Philosophy

in

Economics

Stony Brook University

2016

This dissertation consists of three chapters surrounding the impacts of health information technology systems (HIT) on hospital inpatient outcomes. In an effort to eliminate inefficiencies in the US health care sector, policymakers have made a concerted effort to encourage hospitals and physicians to adopt health information technology systems. In Chapter 2, I construct a unique dataset on health information technology adoption and health outcomes in New York State to conduct a hospital level analysis identifying the impact of adopting HIT on inpatient outcomes (rates of adverse drug events and severity-adjusted mortality). Unlike previous studies, the patient population is not restricted to Medicare patients, but covers all ages and insurance types. After controlling for unobserved hospital quality and endogenous HIT adoption, my results suggest that a hospital's severity-adjusted mortality decreases by 0.3 percentage points. When restricted to the Medicare patients, I find HIT adoption lowers a hospital's severity-adjusted mortality rate by 0.5 percentage points. I find HIT to have no significant effect on the rate of ADEs.

In Chapter 3, I extend the analysis of Chapter 2 to conduct a patient level analysis identifying the impact of adopting HIT on inpatient mortality for pneumonia, COPD, and CHF inpatients. The econometric analysis requires the use of a binary outcome and binary endogenous variable, and presents challenges in estimation. The merits of two popular estimation methods are discussed, the instrumental variables linear probability model and the bivariate

probit, both of which are of interest to applied researchers. After controlling for unobserved hospital quality and endogenous HIT adoption, my results suggest that HIT adoption significantly reduces a patient's likelihood of dying across all three conditions and the effect size grows when patients are restricted to more homogeneous groups.

In Chapter 4, I discuss extensions to the model including the potential impacts of HIT on costs, readmissions, and length of stay for hospital inpatients. The underlying econometric model for cost-side outcomes is likely to be different than the production based approach taken in the previous chapters and deserves special attention. Each of the above chapters and the data used in the analysis are original contributions to this new and growing area of economic research. Chapter 5 concludes the dissertation.

Dedication Page

To Mom, Dad, Meg, and Pat.

Table of Contents

Contents

List of Figures/Tables/Illustrations

List of Figures

List of Tables

Acknowledgements

There are many people who have helped me reach this point without whom, this dissertation would not have been possible. I'd like to take a moment to thank you all because you've all helped create what's written on the following pages, except for the mistakes, which I selfishly take full credit for.

To Mom, Dad, Meg, and Patrick. Thank you for shaping me into the man that I am, without you none of this would be possible. I love you all more than I can say.

Professor Chris Annala, thank you for believing in me and starting me down this road. You've inspired me more than you'll ever know and I'll never forget the lessons you've taught me.

Professor Peter Morgan and Professor Goncalo Monteiro, you believed in me when no one else would, thank you.

To Professor John Rizzo and Professor Mark Montgomery, thank you for showing me what it takes to be a true economic researcher. Thank you for making the most complex ideas appear simple and giving me the tools necessary to be successful as I venture into academia.

To my all my economics classmates, Bruno, Vasileios, Tamara, Xin, Chang, Tao, Arda and especially Maria Sanmartin. Maria, thanks for reassuring me during moments of doubt, being a constant source of support, and always challenging me to be my best. You are destined for great things.

To my research group, Martin, Lorraine, Jonelle, and Pam. You are some of the most brilliant people that I've ever met and I'm so lucky to have met you.

To Yue-Houng, Jamie, Vinnie, Sean, Joanne, Brian, and Melissa. You are the greatest and best friends that I have ever known. Thank you for keeping me sane and seeing me through all of this.

Finally, to my advisor, Professor Debra Dwyer. You are one of the most selfless and wonderful individuals that I have ever met. You saw potential in me, where I never did. You pushed me to academic heights that I never dreamed possible. You did more for me than any mentor ever has. From the bottom of my heart I thank you for all you've done and I will always work to make you proud.

Chapter 1

1 Introduction

As of 2014, the United States (US) spent $3 trillion on health care, a staggering 20% of GDP NCHS (2013). Relative to the rest of the world, no country spends more per capita on health care and only two (Sierra Leone and Liberia) spend more as a percentage of GDP (WorldBank, 2013). Despite this large expenditure, many measures of health outcomes are worse than those of other OECD countries(OECD, 2012) .

Given the sub-par outcome measures relative to the resources spent on health care, the Obama Administration has made a concerted effort to eliminate inefficiencies in the health care system by encouraging the adoption of Health Information Technology (HIT). Proponents of HIT see it as an effective means to reduce medical errors, costs, and improve patient outcomes by limiting medical mistakes. The RAND Institute projected that, in increasing efficiency and improving quality of care, wide-scale HIT adoption could lead to a reduction of $142-$371 billion in health care expenditure over a fifteen year period (Hillestad et al., 2005). With this in mind, the Obama Administration has targeted wide-scale HIT adoption as a policy goal. Speaking on the matter, Mr. Obama said:

> "To improve the quality of our health care while lowering its costs, we will make the immediate investments necessary to ensure that, within five years, all of America's medical records are computerized. This will cut waste, eliminate red tape and reduce the need to repeat expensive medical tests."(Obama, 2008)

In 2009, as a part of the American Recovery and Reinvestment Act (ARRA), the Administration signed the Health Information Technology for Economic and Clinical Health (HITECH) Act into law. The $35 billion act allocated substantial subsidies to encourage physicians and hospitals to adopt "certified" HIT systems that meet "Meaningful Use" criteria.

Despite the allocation of these funds, there is still much uncertainty surrounding the impact of HIT systems on costs and quality health of care. This dissertation consists of three chapters surrounding the impacts of health information technology systems (HIT) on hospital inpatient outcomes. In Chapter 1, I construct a unique dataset on health information technology adoption and health outcomes in New York State to conduct a hospital level analysis identifying the impact of adopting HIT on inpatient outcomes (rates of adverse drug events and severity-adjusted mortality). Unlike previous studies, the patient population is not restricted to Medicare patients, but covers all ages and insurance types. After controlling for unobserved hospital quality and endogenous HIT adoption, my results suggest that a hospital's severity-adjusted mortality decreases by 0.3 percentage points. When restricted to the Medicare patients, I find HIT adoption lowers a hospital's severity-adjusted mortality rate by 0.5 percentage points, or approximately 117 deaths per year. This result is both robust and significant. I find HIT to have no significant effect on the rate of ADEs.

In Chapter 2, I extend the analysis of Chapter 1 to conduct a patient level analysis identifying the impact of adopting HIT on inpatient mortality for pneumonia, COPD, and CHF inpatients. The econometric analysis requires the use of a binary outcome and binary endogenous variable, and presents challenges in estimation. The merits of two popular estimation methods are discussed, the instrumental variables linear probability model and the bivariate probit, both of which are of interest to applied researchers. After controlling for unobserved hospital quality and endogenous HIT adoption, my results suggest that HIT adoption significantly reduces a patient's likelihood of dying across all three conditions and the effect size grows when patients are restricted to more homogeneous groups.

In Chapter 3, I discuss extensions to the model including the potential impacts of HIT on costs, readmissions, and length of stay for hospital inpatients. The underlying econometric model for cost-side outcomes is likely to be different than the production based approach taken in the previous chapters and deserves special attention. Each of the above chapters and the data used in the analysis are original contributions to this new and growing area of economic research. Chapter 4 concludes the dissertation.

Chapter 2

2 Is HIT a Hit? The Impact of Health Information Technology on Hospital Inpatient Outcomes

2.1 Introduction

In this chapter I focus on identifying the impact wide-scale HIT adoption on hospital inpatient outcomes. This Chapter will be divided into six additional sections. In the second section, a broader definition of HIT is given in addition to an overview of the specific components of HIT that will be utilized throughout the dissertation. In the third Section a review of the relevant literature as well as how this Chapter will contribute understanding the impact of HIT on outcomes will be given. The fourth section introduces and describes the consrturction of a novel dataset that will be brought to bear on the analysis. The fifth section presents the econometric framework and gives special attention to the potential endogeneity of a hospital's HIT adoption decision. The sixth Section presents the results and discusses policy implications, while the seventh section concludes.

2.2 What is Health Information Technology?

One goal of the HITECH Act is to establish an interoperable health information sharing network across providers. However, before this network is established, hospitals and physicians offices must first install and correctly utilize the necessary components of HIT. To guide health care providers in the installation of HIT, the HITECH Act segmented adoption into three stages. In each of these stages, providers that meet criteria known as as "meaningful use" requirements, are eligible for HITECH subsidy/incentive payments. Stage 1 of meaningful use focuses on data capturing and sharing (2011-2012), Stage 2 focuses on advanced clinical processes (2014) and Stage 3 focuses on improved outcomes (2016). These stages directly motivate which HIT components are studied in this analysis.

It should be noted that HIT is an umbrella term for technology used in health care that can refer to a variety of components, and serves as a "catchall" term (Dranove et al., 2014, 2015). This makes it challenging to define what is meant by a hospital having installed a HIT system. Even within the literature many studies provide slightly differing definitions of HIT or elect to study certain HIT components over others (Agha, 2014; Dranove et al., 2014, 2015; Miller and Tucker, 2011; McCullough et al., 2013). Given that the time frame of my data ends with the terminal year of the first stage of meaningful use, I will focus on those components most correlated with satisfying the Stage 1 criteria.

The three types of technologies I examine are clinical data repositories (CDRs), clinical decision support systems (CDSS), and computerized practitioner order entry (CPOE) systems. Within a hospital, CDRs create and compile an electronic medical record for the patient. In other words, CDRs house and collect data on patients gathered from various departments, allowing physicians to access a comprehensive record of the patient's health history.[1] A CDSS aids physicians in forming patient specific diagnoses, determining patient specific risk factors, and checks for potential drug interactions to prevent potentially harmful adverse drug events.

Relative to CDR and CDSS, CPOEs are a more advanced, are relatively more costly to install, and require more time and effort on behalf of physicians to utilize correctly. In addition to providing clinical support (much like a CDSS), these systems allow physicians to send orders electronically to the department responsible for completing them (ex: radiology or pharmacy). These functionalities are key to satisfying the stages of meaningful use (Dranove et al., 2014, 2015).

In his seminal work, Kenneth Arrow emphasized that problems of information are at the heart of many issues in health care Arrow (1963). While these problems may never be perfectly eliminated, health information technology is one such avenue that may mitigate the negative effects posed by information problems. At their heart, HIT systems seek to provide better information to health care professionals, allowing them to make better decisions when treating patients, thereby decreasing the likelihood of medical errors.

One area where providing clinicians with better information via HIT adoption could have a sizable impact on, is the reduction of averse drug events (ADEs). ADEs refer to the harm a patient undergoes when prescribed an incorrect drug, or when two (or more) drugs the patient is taking interact in a harmful way. These events are best categorized as low-probability high-cost events, both for the hospital and the patient, yet many are preventable. For example, in 2011 there were 700,000 emergency department visits and 120,000 hospitalizations due to ADEs, resulting in $3.5 billion in extra medi-

[1] HIMSS defines a hospital that has installed a CDR as having installed an electronic medical records system.

cal spending. Additionally, 40% of ADEs occurring in the ambulatory setting are considered preventable (CDC, AHRQ).

A HIT system would help physicians to reduce ADEs in two ways: first, it would increase her information set by giving her the most up to date medication list for the patient; second it would provide warning flags for potential ADEs. Thus HIT has the potential to reduce these events resulting in decreases in costs to the hospital and increases in patient welfare. Some hospitals are realizing returns on their HIT investment, such as DeKalb Medical Center in Georgia which credits its HIT implementations with reducing medication administration errors by 66% Association (2010).

However, there are potential pitfalls that can hinder the realization of the benefits of HIT. Implementing an HIT system is far different from installing commercial over-the-counter software and involves an information technology overhaul of the adopting hospital. The total time required for an EMR system to be implemented effectively, counting the installation procedure as well as the necessary staff training, can be substantial. Some facilities have reported that they need 3-6 months to choose a system, 18-24 months to install a system, and up to a 12 months to troubleshoot a system Association (2010). There is also worry that the timeline of the HITECH Act was too aggressive, placing pressure on hospitals to rapidly adopt HIT without being able to take the time to appropriately select the HIT most appropriate for their facility. Furthermore, as with any new technology, staff training can be quite costly and met with resistance, especially among older physicians, making the benefits of HIT vulnerable to Solow's "productivity paradox". Thus there can be substantial learning-by-doing and training costs, however these are expected to dissipate over time.

2.3 Literature Review

Given that the national emphasis on HIT adoption did not happen until recently, the economic literature is small relative to more well established topics. In what follows, I will lay out the key findings from the papers most directly related to the effectiveness of HIT on improving patient outcomes.

One study demonstrating significant promise for HIT systems, in terms of improving patient outcomes, is Miller and Tucker (2011), specifically with regards to neonatal mortality. Using a twelve year panel on county level electronic medical record adoption rates they examine neonatal mortality rates, and control for endogenous county-level adoption using an instrumental variables approach. They find that a 10% increase in HIT adoption in hospitals at the county level will reduce neonatal mortality rates by 16 deaths per 100,000 live births. Additionally, these gains were larger for African American women, Hispanic women, less educated women, and unmarried women, indicating HIT systems also improve equity by providing the primary care physician with more information over the course of the pregnancy. They also found HIT had no ability to reduce acute onset events (ex: SIDS), but were much more effective in reducing conditions that could be tracked overtime (ex: prematuraty and maternal complications). This suggests that HIT is best suited to treat conditions where information is a key component in the treatment, and where the health of the patient may be monitored over time. While powerful, this study is limited in scope as it only covers the topic of neonatal mortality, thereby limiting the study to a specific subset of the population and subsets of outcomes. Additionally, the impacts of wide scale HIT adoption are not likely to be captured as the terminal year of the data used is 2006, three years prior to the passing of the HITECH Act.

Finding less promising results for the effectiveness of HIT systems, Agha (2014) looks at Medicare Part A and Part B claims data. The specific HIT components studied are CDSS and electronic medical record systems. The data comprise a 20% sample of the US inpatient admissions of 3,900 hospitals over the years 1998-2005, and is limited to those with specific conditions for which being admitted is a good proxy for disease incidence.[2] Using a fixed-effects analysis, she finds HIT adoption has little impact on patient mortality, medical complication rates, adverse drug events, and readmission rates. However, much like Miller and Tucker (2011), this study is constrained to a specific population (Medicare patients) which may not capture the implications of wide scale HIT adoption on the general patient population brought about by the HITECH Act.

McCullough et al. (2013) also examine the ability of HIT to improve the quality of inpatient outcomes, while simultaneously controlling for local labor market complementarities that may make HIT systems more effective (ex: concentration of local IT workers, proximity to other hospitals). The components studied here are CDRs and computerized practitioner order entry (CPOE) systems. Using a 2002-2007 panel of Medicare fee-for-service data to conduct a patient-level difference-in-difference analysis, they find no effect on mean patient outcomes (30-day readmission, 60-day mortality, length of stay) for all but the most severely ill. They do find network effects to have a beneficial effect on patient outcomes (easier exchange of information and access to information technology professionals). However, a hospital is said to have a

[2] AMI, stroke, hip fracture, lung cancer, colon cancer, gastrointestinal hemorrhage, or pneumonia.

HIT system installed if a hospital has basic order entry and electronic record capabilities, both of which have long been in existence and on their own are unlikely to produce substantial benefits. Additionally, like the aforementioned studies, this selects upon a specific patient population and the data pre-dates the HITECH Act.

While the current analysis is restricted to the effects of HIT adoption in New York State, it contributes to the literature in significant ways. First, to my knowledge, this is the first study of its kind to include data spanning the time periods before and after the HITECH Act was passed. This will allow for the Act to serve as a natural instrument when controlling for the possibility of endogenous HIT adoption. Additionally, this has the added benefit of allowing the impacts of the wide-scale HIT adoption brought about by the HITECH Act to be more accurately measured. Secondly, the data contain extremely detailed and rich patient level controls allowing these to be used for risk-adjustment at the hospital level. Third, the data are not restricted to subsets of the population as previous studies have been. Namely, I have access to every inpatient admission that occurred in a New York State hospital over the sample period.

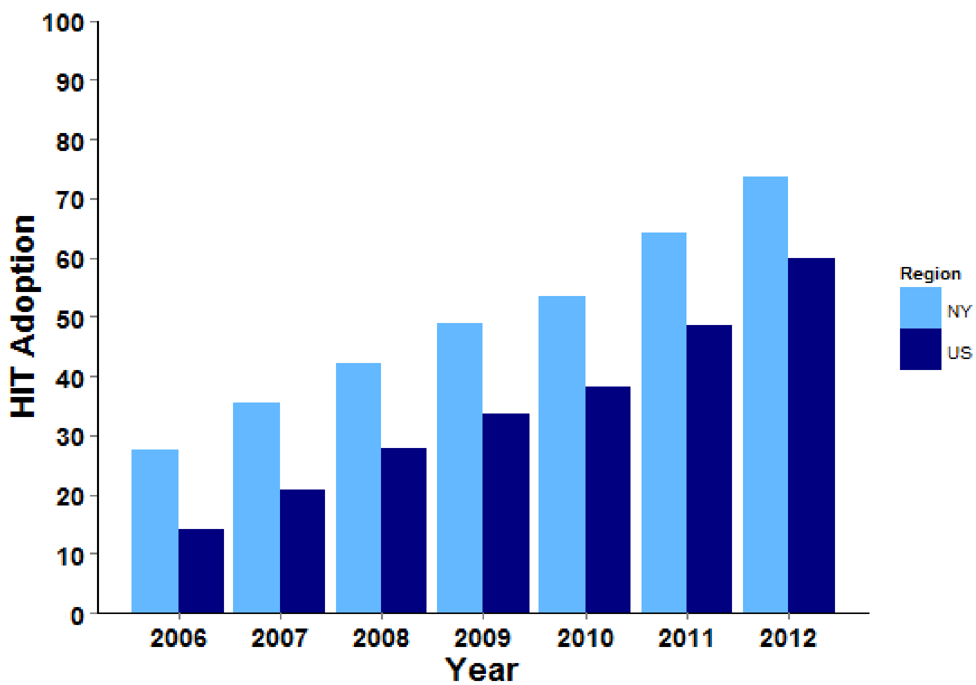


Figure 2.1: National and New York State HIT Adoption

Furthermore, although the sample is limited to New York State, hospitals in New York have had higher rates of HIT adoption than the US national average as seen in Figure 2.1. Given the time costs of installing and learning how to correctly utilize HIT, choosing a state with hospitals that have had HIT components installed over a longer period will better enable this analysis to see the long-run impacts of HIT adoption.

2.4 Data Description

The data come from three sources and compose a panel dataset spanning the years 2006-2012. The first source of data comes from the Statewide Planning and Research Cooperative System (SPARCS), which collects patient level demographic, treatment, insurance, diagnosis, and discharge data for every inpatient admission occurring in a New York State hospital. Unlike Medicare samples used in previous studies, SPARCS tracks data for every inpatient admission in a New York State hospital, not just subsets of the population based on age or insurance/program coverage.[3]

The second source of data comes from the Health Information Management Systems Society's (HIMSS) Analytics Survey, conducted by the Dorenfest Institute. The HIMSS dataset is considered the industry standard for measuring the HIT components that a hospital has adopted. The HIMSS data contain the HIT adoption history and facility information for over 5,300 facilities. The third source of data come from the American Hospital Association's annual survey. This dataset gives additional information on the hospitals that will be used in the analysis (ex: beds, number of doctors, etc.).

An important limitation of the HIMSS data pointed out by (Agha, 2014) is that, while the components a hospital has adopted are visible, given that systems vary from vendor to vendor, the exact capabilities and quality are not observable. Following Agha (2014) and in order to better understand which capabilities I am likely to observe in the HIMSS data, I defer to Jah et al. (2009). In 2008, Jah et al. (2009) conducted a survey of 2952 hospitals and provide insight into which HIT features and capabilities were most common among different HIT components. This survey pre-dates the wide-scale adoption brought about by the HITECH Act but can serve as a glimpse into the capabilities of the systems I observe.

The most common features of CDRs were demographic characteristics, medication lists, discharge summaries, and problem lists. The most common features of CDSS systems were drug-alergy alerts, clinical reminders, and drug-

[3] All patient level characteristics are "rolled up" to hospital level averages.

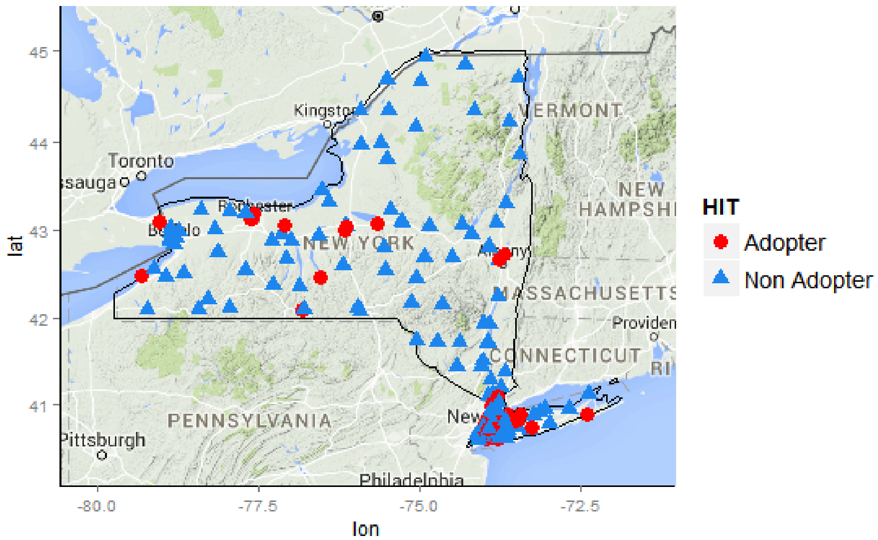
drug interaction alerts. Due to the cost and advanced training necessary to utilize, CPOEs had very low-levels of installation, but the most common features were electonically sending medication lists and laboratory tests. These features are likely to be functionalities of the HIT components I observe, and are all relevant factors for satisfying the meaningful use criteria. With this in mind, I feel comfortable that by examining the adoption of the HIT components under consideration I will be measuring investments in HIT brought about by the HITECH Act.

Merging resulted in a seven-year unbalanced panel with 1,248 hospital-year observations. There is little sample attrition with the majority of changes in the number of hospitals being attributed to closings or mergers.[4] The number and geographical distribution of the hospitals in the sample can be viewed in Figure 2.2. There is substantial variation HIT adotion when comparing the number of adopting hospitals the beginning year of the sample (2006) relative to the terminal year of the sample (2012). This is most likely driven by the incentives provided by the HITECH Act, ultimately accelerating the the take-up of HIT within hospitals.

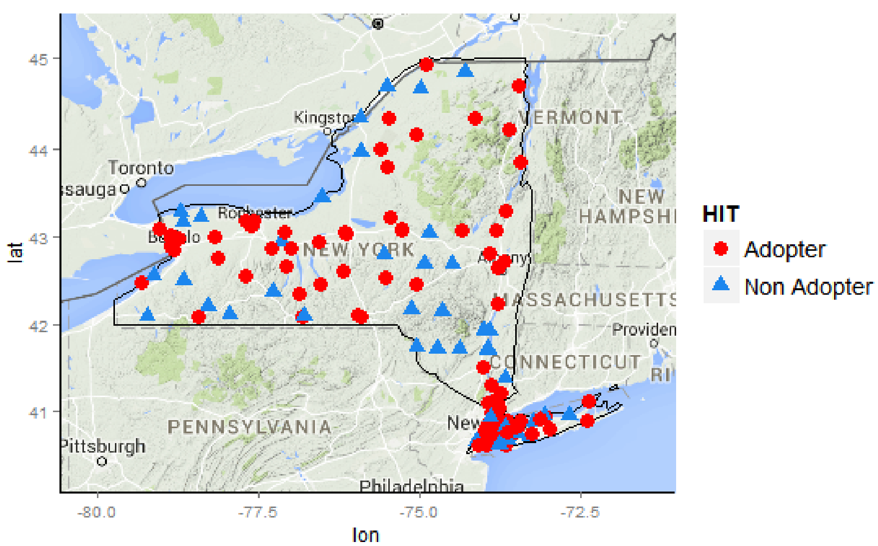
I measure if a hospital has adopted an HIT system by drawing from the meaningful use requirements, the complexity of the HIT component, the HIMSS Adoption Model, and the relevant literature (Agha, 2014; Dranove et al., 2014, 2015; Miller and Tucker, 2011; McCullough et al., 2013). I code a dichotomous HIT variable that takes on a value of one if the hospital has adopted an operational CDR and CPOE or CDSS and CPOE system, and zero otherwise. Both the CDR and CDSS applications are less costly to utilize and install, while CPOEs are more costly in terms of both physician effort and cost. The above definition ensures that "adopters" have at least one "basic" component and one "advanced" component, both of which are necessary for meeting the meaningful use requirements.

The main outcome measures considered in this study are the severity-adjusted mortality rate and the rate of adverse drug events. Mortality rates were calculated at the hospital level and used the patient's discharge status to indicate whether or not the patient had passed away. These rates were then severity-adjusted by using the All Patient Refined Diagnosis Related Group (APR-DRG) weights. These weights take into account the severity of a patient's illness, the patient's risk of mortality, and resource utilization of the patient. Adjusting in this way will allow for an accurate comparison of mortality rates across hospitals.

[4]As an additional check, means tests conclude that there was no statistically significant difference in the pre and post merged data.



(a) 2006



(b) 2012

Figure 2.2: HIT Adopting Hospitals

Adverse drug events (ADEs) are self-reported by hospitals and commonly underreported, making measuring them a challnge. To stay consistent with the literature and ensure comparability of results, I follow Agha (2014) and Hougland et al. (2008) to calculate ADEs. ADEs were calculated using ICD-9 codes that most likely indicate the presence of an adverse event. The specific codes were chosen following the recommendation of an expert panel of medical professionals, and included those ICD-9 codes most likely to represent an ADE. The ADE measurement includes but is not limited to accidental poisonings, incorrect doses, and complications surrounding the use of a prescribed medication. As a further safeguard, "present on admission" indicators included in SPARCS were used to ensure that any ADE measured in the sample happened after the patient had been admitted to the hospital.

Summary statistics are listed in Table 2.1, with the first column of the table showing means across the full sample and the remaining two columns showing statistics contingent upon HIT adoption status. In general, HIT adopters have more full-time employees, doctors, beds, and see more patients than non-adopting hospitals. Additionally, hospitals belonging to a system as well as academic hospitals are more likely to have adopted when compared to the rest of the sample. This is in line with theory as academic hospitals tend to be technology-loving. Additionally, system hospitals tend to adopt together and can be expected to have a greater benefit from HIT, which is realized via information sharing across the system.

2.5 Conceptual Framework

The econometric model implemented in the study is motivated by the general production function framework first developed by Zellner et al. (1966) and adapted to health care by Jensen and Morrisey (1986). I view the selected outcome measures as the end result of a production process undertaken by hospitals. Specifically, I model case-mix adjusted output and health production in hospitals, paying special attention to the HIT capital component and its interaction with labor and other components of the production process. Additionally, this framework informed the selection of the outcome measures that are being studied. Formally, using labor (L), capital (K), demand-side patient attributes (N), and HIT the model is:

$$Y = f(L, K, N, HIT) \tag{1}$$

One criticism of the current literature is the tendency to model all outcome

Variable	**Full Sample**	**Non-Adopter**	**HIT Adopter**
Hospital Characteristics			
HIT	0.4923	NA	NA
Government Owned	0.1314	0.0501	0.2167
Teaching	0.2532	0.1095	0.4039
Joint Commission Accredited	0.8413	0.8138	0.8703
System Member	0.5264	0.4429	0.6141
Full Time Employees	2,127.647	1,193.847	3,107.448
Doctors (Full Time)	67.160	27.682	67.160
Annual Admissions	13,013.530	9,249.951	16,962.500
Beds	367.829	272.685	467.660
Robotic Surgery System	0.2428	0.6870	0.8079
MRI	0.6034	0.5086	0.7028
CT Scanner	0.7492	0.6776	0.7492
Ultrasound Scanner	0.7460	0.6870	0.8079
Patient Characteristics			
Female	0.5702	0.5698	0.5707
White	0.7040	0.7795	0.6247
Black	0.1038	0.1044	0.1885
Other Race	0.1505	0.1160	0.1868
Hispanic	0.0947	0.0837	0.1061
Medicare	0.4144	0.4498	0.3773
Medicaid	0.2257	0.1983	0.2543
Private Insurance	0.3285	0.3229	0.3345
Outcome Measures			
Adverse Drug Events	0.0099	0.0107	0.0090
severity-adjusted Mortality	0.0156	0.0200	0.0109
Count	1248	639	609

[a] All means are reported at the hospital level and weighted by the number of patients.

Table 2.1: Summary Statistics: Full Sample

measures in the same way, following a "kitchen sink" approach. Outcomes such as rates of readmission and length of stay are likely to be driven by cost-side factors and policy measures aimed at reducing cost. Thus, it is likely inappropriate to apply the same model specification and variables as one would for these outcomes, as one would for more traditional outcome measures. The inputs that drive hospital produced patient outcomes, such as mortality, will not be the same as those driving other outcomes that are affected by cost driven policy goals, such as readmission rates. While some work has been done in this area, careful attention should be given to these issues and modeling the impact of HIT on reducing cost should be the focus of future work. Chapter three of this dissertation expands upon this idea.

In the conceptual framework I apply, outcomes are produced at the hospital level. Hospitals produce outcomes by utilizing HIT in conjunction with other inputs of production (beds, hospital staff, doctors, etc.). When utilized appropriately, HIT should work to enhance a clinician's information set thereby allowing them to make more accurate medical decisions. Thus it is appropriate to view HIT as a form of labor augmenting capital, that should allow for the provision of better care and result in the realization of better patient outcomes. Thus, all else equal, the hypothesized effect of HIT on all three outcomes under consideration should be negative.

2.5.1 Econometric Specification

Following from the conceptual framework above, the econometric equation to be estimated for hospital i at time t is:

$$Y_{it} = \alpha + HIT_{it}\beta + \mathbf{X'}_{it}\boldsymbol{\theta} + u_i + \epsilon_{it} \tag{2}$$

Where Y_{it} represents the outcome being analyzed, HIT_{it} is a binary variable that equals one if the hospital has adopted an HIT system and zero otherwise. The remaining covariates $\mathbf{X}_{it}$ represent additional relevant capital, labor, and patient-side attributes used in the production process.

It is very likely that there are unmeasured hospital specific effects that make the OLS estimator inconsistent. For example, some hospitals may have adopted HIT irrespective of HITECH Act's passage due to unobservable hospital level heterogeneity, making the HIT adoption decision endogenous.[5]. Table 2.2 displays summary statistics for the pre-HITECH Act period. It is clear that early adopters of HIT tend to be larger and have access to more resources,

[5]For example adopting hospitals may be more technology-loving, have greater resource availability, or be higher quality.

Variable	**Adopter**	**Non-Adopter**
Admissions	17,779.820	9,858.578
Full Time Employees	3,155.617	1,354.811
Doctors	114.741	30.253
Beds	466.995	308.042
Teaching Hospital	48.756%	13.611%
System Member	66.667%	42.500%
Joint Commission Accreditation	94.030%	85.833%
Mortality	1.097%	1.823%
ADEs	0.873%	1.108%
Number	201	360

[a] The statistics above are for the pre-HITECH Act potion of the sample (2006-2008). All means are reported at the hospital level and weighted by the number of patients.

Table 2.2: Pre-HITECH Act Summary Statistics (2006-2008)

features which are correlated with higher quality. Being a "high quality" may not only be correlated with HIT adoption, but may also be correlated with the likelihood of realizing better outcomes. Additionally, due to their resources and organizational structure, these hospitals may be able to use HIT more effectively than other types of hospitals, in essence having less steep learning curves. If this were the case then the coefficient on HIT adoption would be measuring both the direct and indirect effect of adoption, with the later being driven by unobservables such as quality or technology-loving behavior. Thus unobserved hospital level heterogeneity likely makes the HIT adoption endogenous. It is equally plausible that this effect works in reverse as well. Essentially, the effect of HIT on "poor quality" hospitals, or those lacking the resources to correctly implement (ex: IT staff, resistance to changing hospital culture, or negative attitudes towards technology) would be biasing the effect of HIT in the opposite direction. For this reason, a priori, it is difficult to determine the direction of the bias introduced from unobserved heterogeneity.

I control for endogeneity of the HIT variable in two ways. First, I take first differences of (2) in order to purge any unobserved time invariant hospital specific factors, u_i, that could be correlated with HIT adoption (ex: ownership structure, departmental management structure). At this point, differences within hospitals may be compared across hospitals. Letting "Δ" represent a first-difference transformation, the model is:

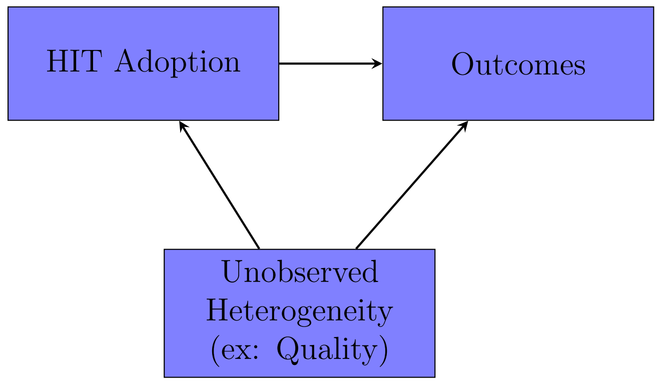
$$\Delta Y_{it} = \Delta HIT_{it}\beta + \Delta \mathbf{X'}_{it}\boldsymbol{\theta} + \Delta\epsilon_{it} \tag{3}$$

$$E(\epsilon_{it} - \epsilon_{it-1}|\mathbf{X}_{it} - \mathbf{X}_{it-1}) = 0 \tag{4}$$

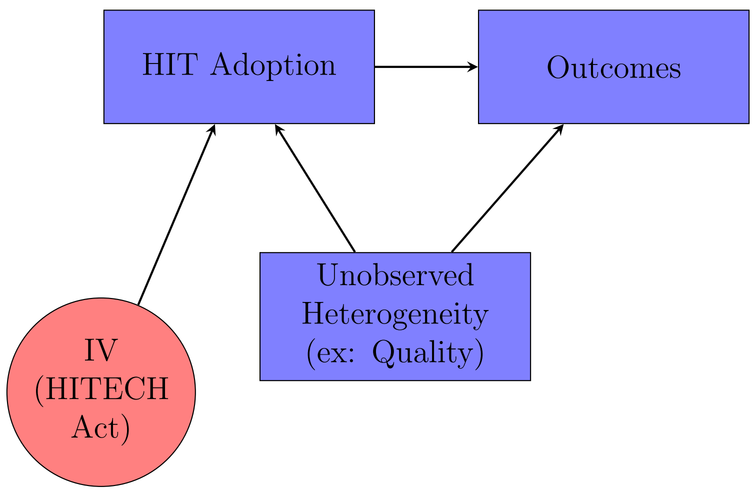
$$E(\epsilon_{it} - \epsilon_{it-1}|HIT_{it} - HIT_{it-1}) \neq 0 \tag{5}$$

Where equation (4) assumes that the auxiliary regressors are weakly exogenous, or that their past realizations are uncorrelated with future disturbances, a necessary step in implementing any standard first-differences transformation. While equation (3) has removed unobserved time invariant effects, there is still the possibility that HIT adoption is serially correlated with the differenced error term $\Delta\epsilon_{it}$. Hence, as shown in (5), the HIT adoption variable would violate the assumption of weak exogeneity. In other words, unmeasured time-varying hospital factors (such as the addition of a new wing, a merger, or change in management structure), could be potentially correlated with HIT adoption. In order to control for any serial correlation, the panel nature of the dataset is exploited in order to implement a difference-GMM dynamic panal data approach akin to that of Arellano and Bond (1991). The basic notion of the instrumental variables strategy is depicted in Figure 2.3.

The difference-GMM approach allows for the introduction of traditional external instruments as well as GMM style lagged instruments. The HITECH Act itself serves as a natural instrumental variable, being correlated with the propensity to adopt HIT, yet being completely uncorrelated with unobserved hospital factors. To capture the passage of the Act, I construct a binary variable that equals one if the year is 2009 or later and zero otherwise. Although the subsidy payments did not begin until 2011 I choose 2009 as HIT systems take time to install and 2009 represented the point in time where new information became available (passage of the Act) and hospital behavior began to change. Given that the HITECH IV is a time trend variable, it is possible that it could be correlated with other events, notably the financial crisis of 2007-2009. To control for this a binary, indicator capturing the time periods of the 2007-2009 recession is included as an additional covariate. Additionally, I introduce a set of GMM style instruments by including lags of exogenous regressors. Lastly, to control for the possibility of heteroskedasticy all standard errors are robust and clustered at the hospital level. This will allow for the following set of moment conditions that will be used in identifying the parameters of interest. Letting k be the number of exogenous regressors and letting



(a) Endogenous Adoption



(b) IV Correction

Figure 2.3: Instrumental Variables Correction

Z represent the HITECH Act instrumental variable, the moment conditions for the i^{th} unit of the panel are:

$$E[Z_{it} \cdot (\Delta Y_{it} - \Delta HIT_{it}\beta - \Delta \mathbf{X'}_{it}\boldsymbol{\theta})] = 0 \tag{6}$$
$$E[\mathbf{X}_{it-1} \cdot (\Delta Y_{it} - \Delta HIT_{it}\beta - \Delta \mathbf{X'}_{it}\boldsymbol{\theta})] = \mathbf{0}_{kx1} \tag{7}$$

Where Equation (7) makes use of the first-lags of the exogenous regressors, but additional lags are available as well.

It should be noted that since the HITECH instrumental variable is a time trend, it could be picking up on factors other than the passage of the ACT, representing a limitation of the current method of analysis. To lend validity to the identification strategy I show robustness to the set of GMM-style instruments chosen. Additionally, as further tests, I rely on the J-statistic as well as the Kleibergen-Paap F statistic to test the validity and strength of my instruments (respectively). While this is a limitation, I feel the IV-GMM strategy is best suited for dealing with the problem of unobserved hospital quality, especially in the presence of time-varying shocks that change hospital HIT adoption behavior (Recession and changing financial climate).

2.6 Results and Analysis

The results of the analysis are listed in Table 2.3, where all standard errors are cluster-robust at the hospital level. The first and second columns represent OLS and first-differences specifications, respectively, and serve as baselines. The OLS specification includes time dummies and naively ignores the endogeneity of HIT adoption. The first-differences specification is empirically identical to a fixed-effects specification and only controls for unobserved hospital attributes impacting HIT adoption that do not vary over time. The third column contains the results for the main difference-GMM framework. The lower half of the table displays tests of overidentification and weak instrumentation for the difference-GMM specification. While intuitively the HITECH Act appears to meet the criteria for a valid instrument, I rely on these tests to further demonstrate that the chosen set of instruments are both sufficiently correlated with the HIT adoption, but uncorrelated with the differenced error term.

The null hypothesis of the J-statistic is that the instruments are jointly valid, with a rejection casting doubt on the validity of the instruments. The Kleibergen-Paap rk F statistic on the excluded instruments is reported as a test of instrument strength and is robust to the presence of heteroskedasticity. The null hypothesis is that there is only weak correlation between the instrument

	$\hat{\beta}$ **OLS**	$\hat{\beta}$ **FD**	$\hat{\beta}$ **GMM**
Mortality	-0.0005	-0.0000	-0.0037*
	(0.0005)	(0.0005)	(0.0021)
ADE	-0.0015	-0.0004	0.0006
	(0.0010)	(0.0006)	(0.0030)

GMM Diagnostic Tests			
		Mortality	ADE
J-Stat P-Val	–	0.5161	0.7108
KP F Stat	–	17.435	17.435
Stock & Yogo 10%	–	10.27	10.27

[a] The reported coefficient in all specifications is that on HIT adoption (having both CDR and CPOE or CDSS and CPOE). Additional controls include: number of doctors, full time employees, the average daily census, beds, teaching hospital status, system hospital status, recession indicator, patient sex, patient race, patient insurance, and patient age (5-year bins). Instruments in the difference GMM specification for mortality and ADEs are: The HITECH Act,the first lag of annual admissions, the first lag of the number of full-time hospital employees, and the interaction of the HITECH Act and the first lag of the proportion of Medicare patients the hospital sees.

[b] *** significance at 1% level, ** significance at 5% level, * significance at 10% level.

Table 2.3: Primary Results

set and the endogenous regressor, with a failure to reject the null casting doubt on the strength of the instrument set. Critical values for this statistic are drawn from Stock and Yogo (2005).[6]

The first row of Table 2.3 shows the results of HIT adoption on a hospital's mortality rate. Both the OLS and first-difference specification indicate that HIT adoption negatively impacts mortality, but are insignificant. Column three applies the difference-GMM specification, making use of the HITECH Act as well as covariates as GMM-style instruments. The GMM specification indicates that HIT negatively impacts mortality and is significant at the 10% level. The results indicate HIT adoption is associated with a 0.37 percentage point reduction in a NYS hospital's inpatient severity-adjusted mortality rate. The diagnostic statistics indicate that the instrument sets are jointly valid and sufficiently strong. To put this result in perspective, data from the National Hospital Discharge Survey indicated that there were 715,000 inpatient deaths in the US in 2010. If every hospital across the US were to realize the benefits found above, this would amount to a reduction of approximately 2,645 deaths across the US. It is of note that the OLS coefficient is relatively smaller when compared to the difference-GMM specification. One potential reason for this is that the HITECH Act incentivized smaller, relatively high-mortality hospitals to adopt when compared to those that had already adopted as Table 2.2 shows. Figure 2.4 corroborates this notion, as the mortality rate of early adopters is declining prior to the passage of the HITECH Act, yet rises after its passage. This increase is due to higher mortality hospitals adopting HIT, and since the benefits of HIT take time to realize, this inflates the mortality rate of adopters. Thus any specification not taking into account the engogeneity of HIT adoption will likely underestimate the impact of HIT adoption.

The next rows of 2.3 show the results for adverse drug events. Both the OLS and first-difference specification indicate that HIT adoption negatively impacts ADEs, but are insignificant. The difference-GMM specifications in column 3 tells much the same story and is also statistically insignificant. This could be a result of the index used to calculate ADEs being too noisy and not accurately identifying ADEs or could simply suggest HIT has not yet impacted the rate of ADES in a meaningful manner.

To test the robustness of the above results, I run the difference-GMM specification using five different sets of GMM-style instruments. The results are shown in Table 2.4, with the combination of instruments used listed below the

[6]The Craig-Donald Wald statistic is more commonly used to test for weak instruments, but is not valid when the iid assumption on the error is dropped. While the Stock and Yogo critical values are valid under the iid assumption Baum and Schaffer (2007), suggest these values as well as the Staiger and Stock (1997) "rule of thumb" of ten may as a guide.

Difference-GMM Robustness Check					
	(1)	(2)	(3)	(4)	(5)
Mortality	-0.0038*	-0.0033	-0.0039*	-0.0029	-0.0028
	(0.0021)	(0.0021)	(0.0021)	(0.0024)	(0.0021)
J-Stat P-Val	0.5161	0.2697	0.4505	0.3971	0.1754
KP F Stat	22.866	17.147	16.293	20.326	14.452
	(1)	(2)	(3)	(4)	(5)
ADEs	0.0007	0.0004	0.0004	-0.0002	0.0015
	(0.0030)	(0.0030)	(0.0030)	(0.0032)	(.0030)
J-Stat P-Val	0.6370	0.7556	0.7449	0.6747	0.2950
KP F Stat	22.866	17.147	16.293	20.326	14.452

[a] All specifications above contain the covatiates listed in the initial model in Table 2.3. Each column contains a different set of lagged GMM-style instruments. Column (1) contains the HITECH IV, first lags of annual admission, and an interaction between the HITECH Act and the lag of the proportion of Medicare patients the hospital sees. Column (2) contains the HITECH IV, first lags of annual admission, an interaction between the HITECH Act, the first lag of the proportion of Medicare patients the hospital sees, and first lag of teaching hospital status. Column (3) contains the HITECH IV, first lags of annual admission, an interaction between the HITECH Act, the first lag of the proportion of Medicare patients the hospital sees, and first lag of number of beds. Column (4) contains the HITECH IV, an interaction between the HITECH Act, the first lag of the number of full time employees. Column (5) contains the HITECH IV, first lags of annual admission, an interaction between the HITECH Act, the lag of the proportion of Medicare patients the hospital sees, first lag of teaching hospital status, and the first lag of belonging to a hospital system.

[b] *** significance at 1% level, ** significance at 5% level, * significance at 10% level.

Table 2.4: Robustness Check

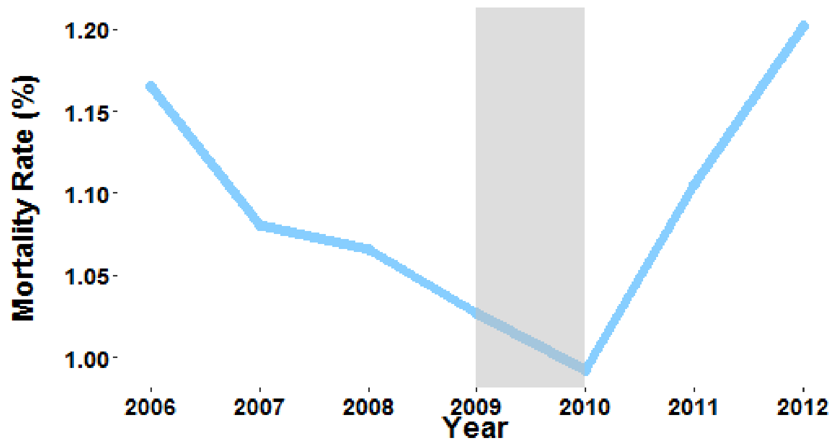


Figure 2.4: HIT Adopter Mortality Rate, Full Sample

table. The top half of Table 2.4 shows that the results for severity-adjusted mortality, fluctuate between a reduction 0.29 to 0.38 percentage points. While all specifications pass the diagnostic tests, lending a measure of credibility to the instruments, the statistical significance fluctuates across the specifications. To further insure that these findings are not being driven by the face that mortality could naturally be falling over time, Figure 2.5 plots the severity-adjusted mortality rate in NYS over the sample period. From this descriptive point of view, there is no obvious trend that mortality is naturally falling, lending support to the findings. However, although the results are not completely robust to the instrument set chosen, that fact that some evidence that HIT is reducing mortality rates for a general patient population is encouraging.

The results on ADEs are less promising, and do not indicate robustness as the magnitude and sign of the coefficient fails to remain consistent across the specifications. This is most likely due to the noise in the ADE index itself, as it is derived from ICD-9 codes, and calls attention for the need of a more accurate measure of these events. It is still possible that HIT is reducing ADEs, as those inpatients that experience an ADE are much more likely to die than their counterparts that do not. Since there seems to be a reduction in mortality, it is plausible that part of this reduction could be driven by a reduction in ADEs.

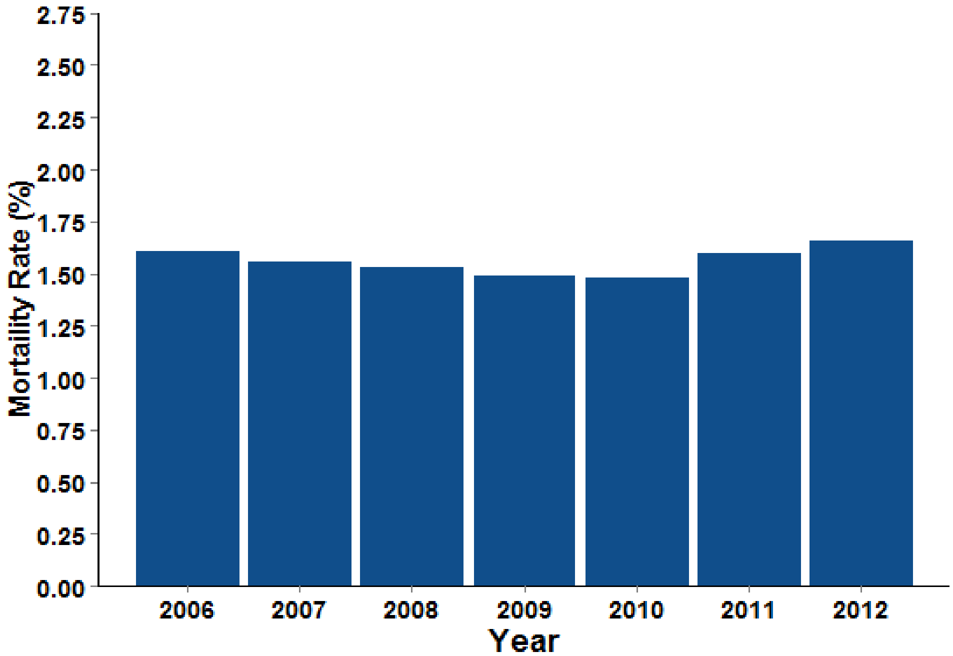


Figure 2.5: severity-adjusted Mortality in NYS

2.6.1 Extension: Medicare Population

One advantage of the data used in this analysis is that it allowed for the impacts of HIT to be examined in a macro-analysis for the general patient population as a whole. While some evidence was found suggesting that HIT adoption could lower mortality, it was dependent upon the instrument set chosen. In this section, I follow the literature and examine the impacts of HIT adoption on Medicare patients, who are generally more ill relative to the patient population as a whole (Agha, 2014; McCullough et al., 2013). When patients are more critically ill, accurate medical information becomes increasingly valuable in order to ensure a positive outcome for the patient. This notion is supported by McCullough et al. (2013) who find that HIT significantly reduces patient mortality for the most severely ill patients in their sample. In this section I seek to discern the impact of HIT adoption on outcomes for Medicare inpatients. Restricting the study to this subsample serves three key purposes: first it will allow us to see if there are larger welfare gains for this segment of the population as opposed to the general inpatient population. These inpatients are generally more ill, with a mean severity-adjusted mortality rate of 2.3% across the sample period. Second the homogeneity of this subsample will help

	$\hat{\beta}$ **OLS**	$\hat{\beta}$ **FD**	$\hat{\beta}$ **GMM**
Mortality	-0.0010	-0.0001	-0.0051***
	(0.0008)	(0.0006)	(0.0019)
ADE	-0.0020	0.0004	0.0027
	(0.0013)	(0.0008)	(0.0029)

GMM Diagnostic Tests			
		Mortality	ADE
J-Stat P-Val	–	0.2459	0.5465
KP F Stat	–	22.986	22.986
Stock & Yogo 10%	–	10.27	10.27

[a] The reported coefficient in all specifications is that on HIT adoption (having both CDR and CPOE or CDSS and CPOE). Additional controls include: number of doctors, full time employees, the average daily census, beds, teaching hospital status, system hospital status, recession indicator, patient sex, patient race, patient insurance, and patient age (5-year bins). Instruments in the difference GMM specification for mortality and ADEs are: The HITECH Act,the first lag of annual admissions, the first lag of the number of full-time hospital employees, and the interaction of the HITECH Act and the first lag of the proportion of Medicare patients the hospital sees.

[b] *** significance at 1% level, ** significance at 5% level, * significance at 10% level.

Table 2.5: Medicare Results

in accurately identifying the impact of HIT. Lastly, it allows for the examination of an extremely policy relevant segment of the population, that will only grow in importance as Medicare enrollment increases.

The results for this section are listed in Table 2.5. The impact of HIT on the severity-adjusted mortality rate for Medicare patients is encouraging. In all three specifications the impact of HIT has the expected sign, with the difference-GMM specification indicating statistical significance at the 1% level. The coefficient indicated that HIT adoption would reduce a hospital's severity-adjusted mortality rate amongst its Medicare inpatients by 0.5 percentage points. If applied the the average number of Medicare deaths in my sample, this would amount to an annual reduction of 117 deaths per year. To test the robustness of the above results I examine how sensitive the results are to different combinations of the instruments. This robustness check is listed in Table 2.6. These results are encouraging as the coefficient on HIT adoption does not vary greatly in magnitude and remains highly significant. While all of the above results passes the diagnostic tests for instrument validity and weakness, these results indicate a substantial effect size and are likely indicative of a ceiling. However, given that NYS hospitals have had these systems longer and may be further along the "learning curve" than other states' hospitals this size is not out of the realm of possibility for such an ill patient population and are overall encouraging.

Turning to the results on the impact of HIT adoption on adverse drug events in Table 2.5, these results largely corroborate those of the previous section. Even within the subsample of Medicare inpatients, no discernible effect of HIT adoption on ADEs is found. This non-result is robust to different sets of GMM-style instruments as is shown in Table 2.6 and corroborates the findings of Agha (2014) who also applied the same ADE measure to Medicare inpatients. It is possible that the ADE index is too noisy to accurately measure ADEs. If this were true, HIT may still be reducing the rate of ADEs as proponents of HIT would anticipate. If HIT was actually reducing ADEs, this reduction could in part be captured by the reduction in morality mentioned above. However, this calls attention to the need of a better measure of adverse events.

2.7 Conclusion

In light of recent policy efforts to encourage hospitals to adopt HIT, this study examines the impacts of HIT adoption in New York State hospitals. To my knowledge this is the first paper studying the time periods before and after the HITECH Act was passed. This allowed for the use of the HITECH Act as a natural experiment in controlling for endogenous HIT Adoption. The results indicate that the impact of HIT on reducing mortality is promising. The results suggest that hospitals that have adopted the either a CDS and CPOE or a CDS and CDSS realize a reduction of 0.3 percentage points in their severity-adjusted mortality rates. This is is the first evidence of HIT having an improvement for a general patient population rather than just a subset of patients. Furthermore, when considering only Medicare inpatients, HIT lowers a hospital's severity-adjusted mortality rate by 0.5 percentage points, or 117 death per year, indicating substantial welfare gains for this group. I found no impact of HIT on a hospital's rate of adverse drug events in any setting. The weakness of these results is most likely driven by a combination of measurement error and hospitals still learning how to fully utilize HIT.

Difference-GMM Robustness Check					
	(1)	(2)	(3)	(4)	(5)
Mortality	-0.0053***	-0.0050***	-0.0062***	-0.0058***	-0.0047***
	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0019)
J-Stat P-Val	0.2176	0.2113	0.4312	0.2600	0.1707
KP F Stat	17.126	19.782	30.764	24.328	19.424
	(1)	(2)	(3)	(4)	(5)
ADEs	0.0027	0.0017	0.0023	0.0022	0.0022
	(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.1687)
J-Stat P-Val	0.1465	0.2425	0.5552	0.5465	0.1687
KP F Stat	17.126	19.782	30.764	24.328	19.424

[a] All specifications above contain the covatiates listed in the initial model in Table 2.5. Each column contains a different set of lagged GMM-style instruments. Column (1) contains the HITECH IV, first lags of annual admissions, first lag of the proportion of Medicaid patients, first lag of beds, and an interaction between the HITECH Act and the lag of the proportion of Medicaid patients the hospital sees. Column (2) contains the HITECH IV, first lags of annual admissions, first lag of the proportion of Medicaid patients, first lag of beds, first lag of teaching status, first lag of belonging to a system, and an interaction between the HITECH Act and the lag of the proportion of Medicaid patients the hospital sees. Column (3) contains the HITECH IV, first lags of annual admissions, and the first lag of the proportion of Medicaid patients. Column (4) contains the HITECH IV, and an interaction between the HITECH Act. Column (5) contains the HITECH IV, first lags of annual admissions, first lag of the proportion of Medicaid patients, first lag of teaching status, and an interaction between the HITECH Act and the lag of the proportion of Medicaid patients the hospital sees.

[b] *** significance at 1% level, ** significance at 5% level, * significance at 10% level.

Table 2.6: Medicare Robustness Check

While the results on mortality are encouraging, there is room for much improvement. A natural extension would be a micro-level analysis examining the role of HIT on diagnoses groups where information plays a key role in the course of care. Additionally, future work should focus on better identifying a hospital's compliance with the meaningful use standards and seeing if results improve in the future, as hospitals progress through the learning curve associated with HIT.

Chapter 3

3 The Impact of HIT on High-Risk Patient Outcomes, a Micro-Level Analysis

3.1 Introduction

In an effort to reduce costs, improve quality, and enhance efficiency within the health care system the US passed the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act. The Act allocated substantial subsidies incentivising the adoption of health information technology (HIT) systems within hospitals and physicians offices. Examining the impacts of this technology at the hospital level, the first chapter of this dissertation addressed the role of health information technology systems on reducing severity-adjusted mortality and adverse drug events. The results indicated that there was suggestive evidence that HIT was reducing mortality for the general patient population, but this result was not robust. When restricted to the Medicare inpatient population, a generally more homogeneous population with more commodities, HIT adoption was shown to significantly reduce a hospital's severity-adjusted mortality rate. This chapter seeks to depart from the macro-level analysis, of the first chapter to a much more granular micro-level analysis. The discharge data introduced in chapter 1 will be utilized to conduct an admissions-level analysis on the impact of HIT on high-risk and high-mortality conditions.

3.2 Literature Review

At their heart, health information technology systems provide clinicians with better information on their patients, ideally leading to improved outcomes of care. I limit the types of HIT studied to those most closely related to the HITECH Act as well as those examined in previous studies. The three types of technologies I examine are clinical data repositories (CDRs), clinical decision support systems (CDSS), and computerized practitioner order entry (CPOE) systems. Within a hospital, CDRs create and compile an electronic medical record for the patient. In other words, CDRs house and collect data on patients gathered from various departments, allowing physicians to access a comprehensive record of the patient's health history.[7] A CDSS aids physicians

[7]HIMSS defines a hospital that has installed a CDR as having installed an electronic medical records system.

in forming patient specific diagnoses, determining patient specific risk factors, and checks for potential drug interactions to prevent potentially harmful adverse drug events.

Realtive to CDR and CDSS, CPOEs are a more advanced, are relatively more costly to install, and require more time and effort on behalf of physicians to utilize correctly. In addition to providing clinical support (much like a CDSS), these systems allow physicians to send orders electronically to the department responsible for completing them (ex: radiology or pharmacy). These functionalities are key to satisfying the stages of meaningful use (Dranove et al., 2014, 2015).

These systems are likely to have the greatest impact on conditions that are chronic in nature. This is especially true for high-mortality conditions, where the role of accurate information is even more important and is often called upon in split-second decisions. Patients suffering from chronic diseases will see more clinicians, likely have more diagnostic testing, and medications, all of which will allow HIT system to collect more information on the patient and screen for potential medication errors. This is in contrast to high mortality acute conditions, such as heart attacks, where HIT systems are unable to collect information before the patient expires. While in principle the error-checking and ability to make patient-specific treatment plans could presumably help in acute cases, the effect is likely much smaller compared to high-mortality chronic conditions (McCullough et al., 2010).

The idea that the impact of HIT is greatest for conditions where information management is paramount is well documented in the literature. In examining the impact of HIT on neonatal mortality, Miller and Tucker (2011) find that a 10% increase in HIT adoption in hospitals at the county level will reduce neonatal mortality rates by 16 deaths per 100,000 live births. They find HIT had no ability to reduce acute onset events (ex: SIDS), but were much more effective in reducing conditions that could be tracked overtime (ex: prematuraty and maternal complications).

Finding less promising results for the effectiveness of HIT systems, Agha (2014) looks at Medicare Part A and Part B claims data and large set of chronic diseases. The specific HIT components studied are CDSS and electronic medical record systems. The data comprise a 20% sample of the US inpatient admissions of 3,900 hospitals over the years 1998-2005, and is limited to those with specific conditions for which being admitted is a good proxy for disease incidence.[8] Using a fixed-effects analysis, she finds HIT adoption has

[8] AMI, stroke, hip fracture, lung cancer, colon cancer, gastrointestinal hemorrhage, or pneumonia.

little impact on patient mortality, medical complication rates, adverse drug events, and readmission rates.

The study most closely related to this study is McCullough et al. (2013). Using a 2002-2007 panel of Medicare fee-for-service data to conduct a patient-level difference-in-difference analysis, they find no effect on mean patient outcomes (30-day readmission, 60-day mortality, length of stay) for all but the most severely ill. The conditions under consideration are BLANK.

This study seeks to improve upon the previous studies in several ways. First, this study directly controls for the fact that HIT adoption is endogenous by using the HITECH Act as a natural instrument. Secondly, the data utilized are extremely detailed and contain every inpatient admission in a New York State Hospital from 2006-2012, and are this not limited by payer-type. Third the time period under consideration covers period before and after the HITECH Act was passed, and thus captures the wide-scale technology adoption spurred about by the Act. Lastly, this study compares the outcomes for two popular binary outcomes estimation methods that retain consistency in the presence of an endogenous binary regressor.

The conditions that will be examined in this study are pneumonia, coronary artery disease (CAD), chronic obstructive pulmonary disorder (COPD), and congestive heart failure (CHF). These conditions are all high-mortality, have been examined in previous HIT studies, and are of policy relevance, notably for the Medicare population. Pneumonia is an infection of the lungs, leading to about 1,000,000 hospital admissions and 50,000 deaths within the United States. The highest risk groups are children, smokers, and those 65 years of age or older. The average length of stay is 5.4 days and often involves care coordination across many health professionals, making information management and communication a key component of treatment(Huang SS, 2011).

COPD is a group of respiratory conditions made up primarily of chronic bronchitis and emphysema and is not completely reversible. In 2008, there were about 822,500 inpatient stays for COPD across the US with an average length of stay of 4.8 days and primarily affects those 65 years and older. The chronic nature of the disease makes information management a crucial aspect in the delivery of care (Wier, 2011).

CHF is a condition where the heart can no longer pump enough blood to sustain and support other organs in the body. It does not necessarily mean that the heart is no longer functioning, but is an extremely serious disease with over half of those afflicted dying within 5 years of being diagnosed. Patients with CHF often readmitted to the hospital several times (50% revisit within 6 months), are subject to many diagnostic tests (MRI, EKG, ECG, stress tests, etc.), and are often prescribed many drugs (beta-blockers, ACE inhibitors, etc.), making information management a crucial component of care(Desai and Stevenson, 2012; Hall MJ, 2012).

Thus all three conditions are chronic in nature and high mortality, making information management an important component in the process of care. Thus, the hypothesized impact of HIT will be to reduce the probability of dying for pneumonia, COPD, or CHF inpatients.

3.3 Data Description

The data come from three sources and compose a panel dataset spanning the years 2006-2012. The first source of data comes from the Statewide Planning and Research Cooperative System (SPARCS), which collects patient level demographic, treatment, insurance, diagnosis, and discharge data for every inpatient admission occurring in a New York State hospital. Unlike Medicare samples used in previous studies, SPARCS tracks data for every inpatient admission in a New York State hospital, not just subsets of the population based on age or insurance/program coverage. Clinical classification diagnosis (CCS) categories were used to identify patients with the conditions under consideration.[9]

The second source of data comes from the Health Information Management Systems Society's (HIMSS) Analytics Survey, conducted by the Dorenfest Institute. The HIMSS dataset is considered the industry standard for measuring the HIT components that a hospital has adopted. The HIMSS data contain the HIT adoption history and facility information for over 5,300 facilities.

The third source of data come from the American Hospital Association's annual survey. This dataset gives additional information on the hospitals that will be used in the analysis (ex: beds, number of doctors, etc.).

I measure if a hospital has adopted an HIT system by drawing from the meaningful use requirements, the complexity of the HIT component, the HIMSS Adoption Model, and the relevant literature (Agha, 2014; Dranove et al., 2014, 2015; Miller and Tucker, 2011; McCullough et al., 2013). I code a dichotomous HIT variable that takes on a value of one if the hospital has adopted an operational CDR and CPOE or CDSS and CPOE system, and zero otherwise. Both the CDR and CDSS applications are less costly to utilize and install, while CPOEs are more costly in terms of both physician effort and cost. The above definition ensures that "adopters" have at least one "ba-

[9]CCS classifications were developed by the Agency for Healthcare Research and Quality (AHRQ) in order to readily identify patients with certain diseases and conditions. There are currently 285 mutually exclusive categories

Variable	Non-Adopter	Adopter
Robotic Surgery	13.615%	35.468%
Ultrasound	68.701%	80.788%
MRI	50.861%	70.280%
CT	67.762%	82.430%
Teaching	10.955%	40.394%
System	44.288%	61.412%
Critical Access	10.485%	2.299%
Doctors (FT)	27.682	108.583
Beds	272.685	467.660
Number	639	609

Table 3.1: Full Sample, HIT Adopters vs Non-Adopters

sic" component and one "advanced" component, both of which are needed for meeting the meaningful use requirements.

Hospital adoption of HIT can be seen in Table 3.1. It is clear that HIT adopters and non-adopters differ in observable ways. Adopters tend to be larger (more beds, doctors, full time employees) and are more likely to have complementary technology systems installed (CT, MRI, ultrasound, and robotic surgery). Adopters are also more likely to be a part of a system (to which information sharing across system members may yield larger returns), are more likely to be a teaching hospital (shown to be technology-loving), and less likely to be a critical access hospital. While there are clear observable differences, there are likely unobserved differences that will drive HIT adoption. For instance unobservables (IT resources, management structure, physician's attitudes) will all impact outcomes and will also impact the propensity of a hospital to adopt HIT. Thus the endogeneity of adoption must be controlled for and will be discussed in at greater length in the econometric specification.

The main outcome measure considered in this study is mortality. Mortality of the patient is measured at the time of discharge, resulting in a binary indicator for mortality.

After merging, there were 375,717 pneumonia admissions, 219,105 COPD admissions, and 398,263 CHF admissions. The descriptive statistics of describing these admissions are in Table 3.2. The majority of the inpatient population across these diseases tends to be older, white, and have Medicare as their primary payer. The mortality rates for these conditions tend to be high relative

Variable	**Pneumonia**	**COPD**	**CHF**
Age	60.956	70.333	73.974
Mortality Rate	4.543%	2.394%	4.289%
White	67.823%	76.973%	64.998%
Black	15.119%	12.170%	20.525%
Hispanic	11.305%	7.108%	9.187%
Medicare	57.973%	71.965%	75.423%
Medicaid	18.470%	11.853%	10.897%
Private Insurance	22.158%	14.803%	12.564%
Other Insurance	1.399%	1.379%	1.117%
Homeless	4.352%	3.790%	3.902%
Foreign Resident	0.478%	0.400%	0.371%
Revisited a Hospital	16.520%	41.052%	42.759%
Minor Mortality Risk	35.846%	35.090%	13.149%
Moderate Mortality Risk	38.002%	37.156%	46.398%
Major Mortality Risk	19.503%	22.685%	30.043%
Extreme Mortality Risk	6.649%	5.070%	10.410%
Minor Severity of Illness	15.140%	17.498%	9.182%
Moderate Severity of Illness	41.569%	42.462%	41.714%
Major Severity of Illness	34.072%	33.758%	40.584%
Extreme Severity of Illness	9.219%	6.281%	8.519%
CMI	1.234	1.106	1.709
Count	375717	219105	398263

[a] The severity of illness (SOI) and risk of mortality (ROM) are derived data elements utilizing information about the patient's comorbidieites and demographic information. These indicators are assigned to the discharge using the All Patient Refined Diagnostic Related Grouper software, developed by 3M Health Information Systems. This information is used to categorize patient records for reimbursement and research purposes.

[b] Other Insurance contains: no insurance, no charge, reimbursement from a correctional facility, and "other".

Table 3.2: Patient Characteristics by Condition

to others ranging from 2.4% - 4.5%. These patients also tend to be relatively ill as fewer than 20% of those admitted had a illness severity that could be classified as "minor".

3.4 Econometric Specification

The underlying model will follow a latent variables framework. Letting Y^* represent a latent indicator of illness, Y to be the outcome observed to the researcher, $\mathbf{X}$ to represent relevant demand-side and hospital covariates, and HIT to represent a binary indicator for an HIT system being installed, the model is:

$$Y^* = \mathbf{X}'\boldsymbol{\beta} + HIT\theta + \epsilon \tag{8}$$

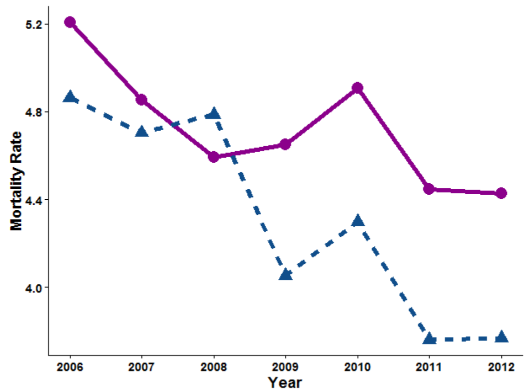
$$Y = \begin{cases} 1, \; if \; Y^* > 0 = \; \epsilon > -(\mathbf{X}'\boldsymbol{\beta} + HIT\theta) \\ 0, if \; Y^* \leq 0 = \epsilon \leq -(\mathbf{X}'\boldsymbol{\beta} + HIT\theta) \end{cases} \tag{9}$$

The actual severity of illness of the individual Y^* is unobserved, it is only when it crosses some threshold does the patient expire and the researcher observe the outcome of death ($Y = 1$). If this threshold is not crossed, then the researcher observes that the patient lives ($Y = 0$). Given that my measure of mortality is binary I will be drawing from a family of binary estimators.

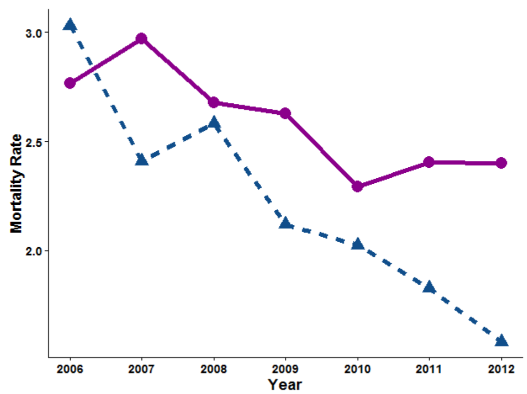
One challenge of the current study is that HIT adoption is endogenous. For instance a poor quality hospital may not adequately utilize HIT (ex: not train staff properly) and will also realize worse outcomes. Conversely good quality hospital may be both more likely to realize good outcomes and more likely to adopt HIT, and thus overstate the true impact of HIT on reducing mortality. The mortality rates for adopters and non-adopters can be seen in Figure 3.1.

Across all three conditions, HIT adopters have much lower mortality rates than non-adopters. However, there is no way of telling a priori if these rates are biased upwards or downwards based on unobserved quality differences across hospitals. Thus without controlling for this unobserved hospital "quality" component biased estimates of θ in Equation (1) will result. To solve this problem, we instrument HIT adoption by using the HITECH Act of 2009 as a natural experiment. The HITECH Act is correlated with the propensity to adopt HIT, yet uncorrelated with unobserved hospital factors affecting HIT adoption. To capture the passage of the Act, I construct a binary variable that equals one if the year is 2009 or later and zero otherwise.

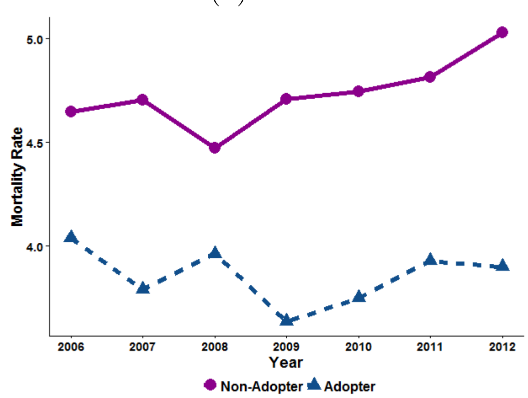
The second challenge comes from the model setup itself. The current situation is that of a binary outcome (mortality) with a binary endogenous regressor



(a) Pneumonia



(b) COPD



(c) CHF

Figure 3.1: Mortality Rates by Condition and HIT Adoption

(HIT adoption) and under this situation, typical estimation methods may not be appropriate. For example control function estimators are generally only consistent only when the endogenous regressor is continuously distributed, if not the latent error term cannot be estimated (Bontemps and Nauges, 2015; Lewbel et al., 2012). The two methods that will be employed in this paper will be an instrumental variables (IV) linear probability model and a bivariate probit. Both of these methods, when given a valid set of instruments, are appropriate for estimating binary outcomes with a binary endogenous regressor and the merits of each will be discussed below (Richard C. Chiburis, 2012).

3.4.1 IV Linear Probability

Given a valid instrument set **Z** the IV linear probability model can be characterized as follows:

$$Y = \mathbf{X}'\boldsymbol{\beta} + HIT\theta + \epsilon \tag{10}$$

$$E[\mathbf{X}\epsilon] = 0,\ E[HIT\epsilon] \neq 0,\ E[\mathbf{Z}\epsilon] = 0 \tag{11}$$

This model is not without its drawbacks, notably the fitted distribution function is assumed to be linear and can easily generate predicted probabilities outside of the unit interval. However, there are many advantages of the model as well. Primarily, these models are computationally simple and can handle general forms of heteroskedasticity, making them a popular choice for applied work (McCullough et al., 2013). No distributional or modeling assumptions need to be made on the endogenous regressor, only the criteria in (4) need to be met. Secondly, despite generating probabilities outside of the unit interval, the main interest of the estimation is the marginal effects. The constant marginal effects offer a simple alternative that are not substantively different from nonlinear methods which often impose strong assumptions on the error (Angrist, 2001).

3.4.2 Maximum Likelihood (Bivariate Probit)

An alternative estimation method to the IV linear probability that can consistently estimate a binary choice model with a binary endogenous variable is the bivariate probit model. This involves the joint estimation of two probit models Letting $I(\cdot)$ denote the indicator function, which takes a value of one when the argument is true and zero otherwise, the model is:

$$Y = I\left(\mathbf{X}\boldsymbol{\beta} + HIT\theta + \epsilon \geq 0\right) \tag{12}$$
$$HIT = I\left(\mathbf{R}'\boldsymbol{\gamma} + \mathbf{Z}'\lambda + \omega \geq 0\right) \tag{13}$$
$$\begin{pmatrix} \epsilon \\ \omega \end{pmatrix} \Bigg| Z \sim \left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right) \tag{14}$$

The bivariate probit model is efficient (when specified correctly), can handle binary endogenous regressors, and allows for heteroskedasticity. However, all of these advantages come at the cost of assuming joint normality of the errors (ϵ, ω) and fully parameterizing the model. While the bivariate probit allows for the calculation of several conditional probabilities, the marginal effect of interest is the unconditional change on mortality when HIT adoption changes Greene (1996).

3.5 Results

While the efficiency of the estimators cannot be compared because the IV linear probability model is incompatible with the threshold crossing model the bivariate probit takes, comparing the magnitude of the estimated effects is still meaningful. As is usual in the case of binary outcomes estimation, the magnitude of the estimated coefficients isn't meaningful, only the sign. To get a better sense of the impact HIT adoption has, I report the estimated marginal effects from both the bivariate probit and IV linear probability models in Table 3.3. For the bivariate probit model, likelihood-ratio tests that the two probit models were unrelated were rejected at the 1% level confirming the endogeneity of HIT adoption.[10]

Across conditions and specifications, the demographic characteristics behave as expected. For example, aging increases the likelihood of dying, while being female, black, and Hispanic decrease the likelihood of dying. Furthermore, the readmission status of the patient was also controlled for. The analysis was conducted at the patient-admission level, so whether or not the patient was readmitted to the hospital was included as a covariate. Clearly, if the patient was readmitted, they have survived at least one previous admission which will decrease the probability of dying. In all specifications, the readmission variable is negative and highly significant. Additionally, the average case-mix index (CMI) of the hospital was included as a regressor to capture the average illness severity of the patient population at that facility. A patient being

[10]This was true for all three conditions under examination.

Variable	**Pneumonia**		**COPD**		**CHF**	
	LPM	BVP	LPM	BVP	LPM	BVP
HIT	-0.0406***	-0.0363***	-0.0442***	-0.0334***	-0.0135	-0.0301***
	(0.0064)	(0.0054)	(0.0066)	(0.0068)	(0.0090)	(0.0073)
Age	0.0010***	0.0018***	0.0009***	0.0012***	0.0014***	0.0016***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
Female	-0.0040***	-0.0064***	-0.0023***	-0.0032***	-0.0033***	-0.0037***
	(0.0070)	(0.0007)	(0.0007)	(0.0008)	(0.0007)	(0.0007)
Black	-0.0034***	0.0004	-0.0050***	-0.0060***	-0.0131***	-0.0149***
	(0.0011)	(0.0012)	(0.0010)	(0.0013)	(0.0009)	(0.0010)
Hispanic	-0.0079***	-0.0109***	-0.0047***	-0.0056***	-0.0076***	-0.0084***
	(0.0010)	(0.0015)	(0.0013)	(0.0017)	(0.0011)	(0.0014)
Medicaid	-0.0033***	0.0002	0.0024**	-0.0019	0.0074***	0.0003
	(0.0011)	(0.0016)	(0.0009)	(0.0017)	(0.0010)	(0.0016)
Private Insurance	-0.0060***	0.0031***	0.0075***	0.0087***	0.0124***	0.0125
	(0.0010)	(0.0012)	(0.0010)	(0.0012)	(0.0011)	(0.0012)
Other Insurance	0.0090***	0.0152***	0.0221***	0.0215***	0.0308***	0.0284***
	(0.0029)	(0.0032)	(0.0035)	(0.0030)	(0.0038)	(0.0030)
Homeless	0.0149***	0.0123***	0.0077***	-0.0022	-0.0064*	-0.0045*
	(0.0027)	(0.0026)	(0.0028)	(0.0029)	(0.0034)	(0.0028)
CMI	0.0432***	0.0437***	0.0245***	0.0234***	0.0109***	0.0115***
	(0.0024)	(0.0023)	(0.0024)	(0.0025)	(0.0009)	(0.0009)
Revisit	-0.0117***	-0.0106***	-0.0033***	-0.0030***	-0.0125***	-0.0127***
	(0.0009)	(0.0010)	(0.0007)	(0.0008)	(0.0006)	(0.0007)

[a] *** significance at 1% level, ** significance at 5% level, * significance at 10% level.

Table 3.3: Marginal Effects, Full Sample

admitted to Hospitals with more severe case mix indicies were more likely to die in all specifications.

To a large degree, socioeconomic indicators of the patient behaved as expected. Whether or not the patient was homeless was included and for all specifications, was positive and highly significant, indicating an increased likelihood of death. Continuing in this vein, "other insurance" (which includes the uninsured), strongly increases the likelihood of dying and is strongly significant.[11] Having Medicaid or a private payer as a primary insurer were not very informative, as the magnitude and significance changed depending on the condition and specification chosen.

For pneumonia, after controlling for endogenous adoption, both models predict that HIT adoption is associated with a reduction in the likelihood of dying and both results are highly significant. The magnitude of the reduction in the likelihood of death across the two models is very similar, and the most similar when compared to the other conditions. The linear probability specification indicates a HIT adoptions causes a reduction in the likelihood of 4.1% while the bivariate probit indicates a reduction of 3.6%. Given the propensity for care coordination and information management required for treating pneumonia, this result is not unexpected and agrees with previous findings (McCullough et al., 2013).

The impact of HIT on reducing mortality in COPD patients appears promising as well. The two specifications yield slightly differing results, both of which are negative and highly significant. On average, the IV linear probability predicts HIT adoption results in a decrease in the likelihood of dying of 4.4% while the bivariate probit predicts a 3.3% reduction. While there is a difference in magnitude, the overall results are encouraging that both specifications identify a significant reduction in the likelihood of dying.

When turning to CHF, there is disagreement among the model results. The linear probability model estimates a reduction in the likelihood of death by 1.4%, yet this is not significantly different from zero nor is it very large in magnitude. The bivariate probit marginal effect is much larger and highly significant at 3%. The disagreement in the results could come from the clinical nature of CHF, and there are arguments to be made both ways. A typical CHF patient may revisit the hospital several times, allowing HIT to accrue more and better information over time. However when the heart finally fails, there may be little that better information can do to prevent the patient from dying. However, it is likely that this disagreement in effect size and magnitude can be attributed to and underlies the importance of selecting the correct model

[11] Medicare was the excluded insurance category.

specification. Apriori, there is no way to tell which model specification is "correct" and selecting the wrong specification can lead to very different results and should serve as a warning to applied researchers. Such disagreement motivates the use of nonparametric and semi-nonparametric methods discussed in Lewbel et al. (2012), which should be the focus of future work.

3.5.1 Extension of Results

One limitation of the present analysis is that there patients who revisit a hospital have already survived a previous visit and thus the errors across the visits will be correlated. In essence, there is a selection effect that makes revisiting patients fundamentally different than patients who do not revisit. Patients that revisit the hospital are different than those who do not revisit, in that they are more likely to live, as Figure 3.2 shows. While this is a limitation of the current analysis, I now restrict my sample to those who do not revisit as one method for eliminating this bias. A more generalized solution to the problem will be left for future work.

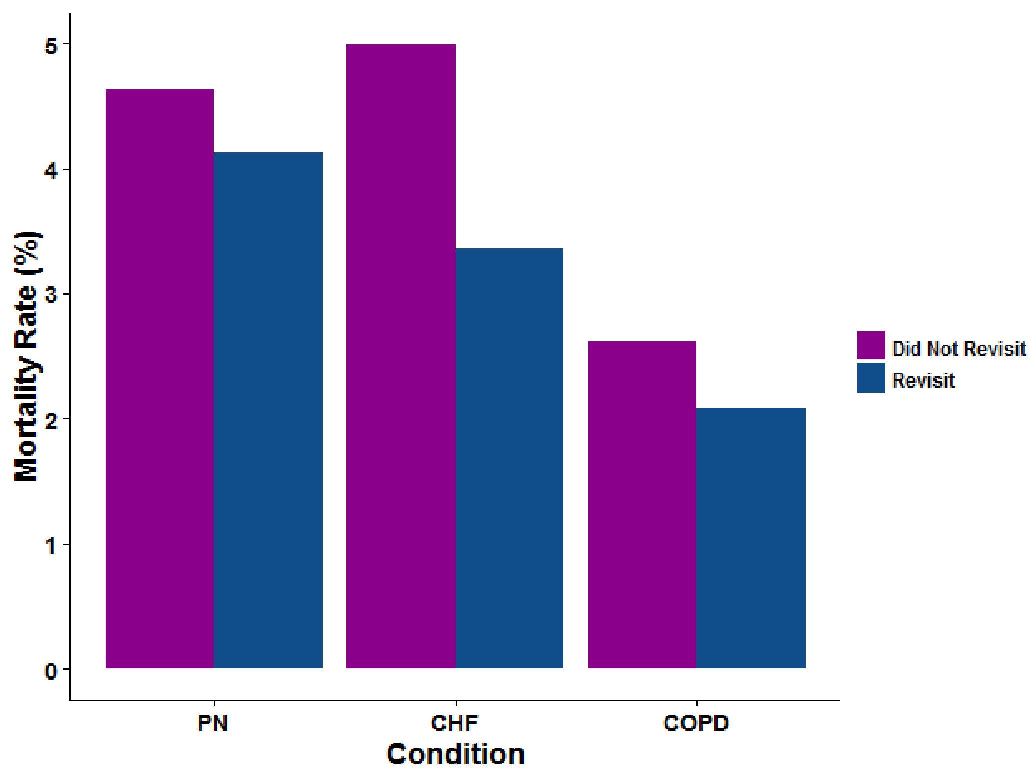


Figure 3.2: Average Mortality by Readmission Status

Table 3.4 shows summary statistics for the included and excluded groups. As expected, the mortality rates for those who do revisit a hospital are lower across the three conditions, with the CHF patients having the largest difference in observed mortality. While there are some differences in the demographics among the patients, for example revisiting pneumonia patients are more likely to have Medicare, the two groups do not appear substantively different aside from the propensity to revisit. On average there is not much difference between the severity of mortality or severity of illness groupings across the conditions. The exception however, is for pneumonia patients where the revisiting population tend to belong to more severe groups on average. For moderately severe, chronic cases, better information and coordination of care could play a large role in securing a positive outcome.

The results marginal effects for the subgroup of patients that do not revisit are listed in Table 3.5. Across all three specifications, the results are highly significant and larger in magnitude than the previous findings. The effect sizes across the two model specifications are quite similar, with the largest difference being 1.1% between the estimated effects for COPD patients. It makes sense that when restricted to a more homogeneous subgroup, that is free from any bias the revisiting behavior may impose, the results improve. Encouragingly, the impact of HIT adoption on the likelihood of dying among CHF patients is now significant and similar in magnitude in both the linear probability and bivariate probit model(5.9%, 6.4% respectively). The improvement in the results and the general agreement in the size of the marginal effects is likely due in large part to the increased homogeneity of the subsample.

3.6 Conclusion

In this study I examine the impact of health information technology adoption on reducing mortality for pneumonia, COPD, and CHF inpatients. In doing so, several contributions are made. The first, deals primarily with the model specification and issues of consistent estimation. Namely, HIT adoption is endogenous, making the analysis at hand one with a binary outcome (mortality) and a binary endogenous variable (HIT adoption). This not often discussed situation is a scenario which renders some popular estimation techniques inconsistent (ex: control functions). The issue of endogneity is addressed by utilizing the HITECH Act of 2009 as a natural experiment for HIT adoption. Additionally the merits and drawbacks of two estimation methods for binary outcomes models with binary endogenous variables are discussed, both of which are of interest to applied researchers. However this is a non-exhaustive summary and newer nonparametric and semiparametric methods

Variable	**Pneumonia**		**COPD**		**CHF**	
Revisit	NO	YES	NO	YES	NO	YES
Age	59.843	66.577	70.462	70.149	74.676	73.035
Mortality Rate	4.625%	4.129%	2.610%	2.083%	4.985%	3.357%
White	66.668%	73.654%	76.719%	77.337%	68.707%	60.032%
Black	15.637	12.502%	11.934%	12.510%	17.396%	24.714%
Hispanic	11.501%	10.313%	7.225%	6.941%	8.090%	10.655%
Medicare	55.741%	69.254%	70.435%	74.161%	75.629%	75.148%
Medicaid	19.073%	15.420%	11.068%	12.980%	9.359%	12.955%
Private Insurance	23.740%	14.163%	17.109%	11.491%	13.858%	10.830%
Other Insurance	1.446%	1.163%	1.387%	1.367%	1.154%	1.067%
Homeless	4.657%	2.811%	3.359%	4.408%	3.636%	4.257%
Foreign Resident	0.431%	0.715%	0.318%	0.517%	0.279%	0.493%
Minor Mortality Risk	38.500%	22.460%	36.971%	32.388%	13.949%	12.078%
Moderate Mortality Risk	36.639%	44.891%	36.183%	38.553%	45.944%	47.007%
Major Mortality Risk	18.335%	25.407%	21.622%	24.210%	29.269%	31.080%
Extreme Mortality Risk	6.526%	7.266%	5.224%	4.850%	10.839%	9.835%
Minor Severity of Illness	16.995%	5.771%	19.314%	14.892%	9.658%	8.544%
Moderate Severity of Illness	42.217%	38.297%	42.560%	42.320%	40.750%	43.005%
Major Severity of Illness	31.948%	44.801%	31.934%	36.378%	40.689%	40.443%
Extreme Severity of Illness	8.840%	11.131%	6.192%	6.409%	8.902%	8.008%
CMI	1.233	1.237	1.106	1.106	1.694	1.730
Count	313647	62070	129158	89947	227968	170295

Table 3.4: Patient Characteristics by Revisiting Behavior

Variable	LPM	BVP
Pneumonia	-0.0670***	-0.0565***
	(0.0076)	(0.0076)
COPD	-0.0588***	-0.0703***
	(0.0086)	(0.0117)
CHF	-0.0592***	-0.0635***
	(0.0112)	(0.0095)

[a] *** significance at 1% level, ** significance at 5% level, * significance at 10% level.

Table 3.5: Marginal Effects, Restricted Sample

such as Lewbel et al. (2012) might be more appropriate.

In addition, I bring to bear an extremely rich patient-level data that contains the entire New York State hospital inpatient population for the years 2006-2012. Thus the analysis is not limited to certain segments of the population or payer type, allowing me to account for important socioeconomic and demographic differences among patients.

Overall I find that HIT adoption significantly reduces the likelihood of dying for patients across all three conditions studied, and this effect grows stronger when the patient population is more homogeneous. This is encouraging and lends support to the mix-findings around the impact of HIT adoption that it is beginning to save lives. However, as systems become more interoperable, health information exchanges improve (allowing for easier information exchange across hospitals), and clinicians become better acclimated with HIT systems, this impact should continue to rise.

Chapter 4

4 Extensions: The Impact of HIT on Costs, Readmissions, and Length of Stay

4.1 Introduction

The HITECH Act of 2009 was passed in an effort to improve efficiency of medical care, improve patient outcomes, and decrease costs. The first two chapters of this dissertation focused primarily on patient outcomes. However, the $35 billion Act represented a substantial outlay of public funds, so it is natural to wonder to what extent HIT will reduce medical costs through efficiency gains. This chapter will discuss how HIT can be utilized to reduce costs and outcomes closely tied to cost-side factors. The objective of this chapter is to lay the groundwork for a more extensive analysis in the future.

With health care costs surpassing the $3 trillion mark, HIT systems were viewed as one way to control costs. In fact Buntin and Cutler (2009) view HIT as a main component of their "Two Trillion Dollar Solution" to modernizing health care. One might reasonably expect cost HIT systems to result in cost-savings due to their ability to coordinate care, reduce duplicate tests, lower lengths of stay, and prevent medical errors. However, it is also possible that these systems could actually result in a temporary increase in costs as they have high these systems take time to learn. Essentially, hospitals and physicians offices could fall victim to Solow's "productivity paradox" whereby new technology causes an initial slowdown until efficiency gains are eventually realized. Determining the true impact of HIT on outcomes such as costs and outcomes linked to costs will be of paramount importance in assessing the success of the HITECH Act in meeting its stayed goals.

4.2 Literature Review

The earliest projections of these cost savings were quite substantial, at $81 billion annually and with a net savings of $371 billion over 15 years in hospitals alone (Hillestad et al., 2005). However, these initial estimated were likely overstated and have been viewed critically in the literature (Sidorov, 2006).

Current reviews on the ability of HIT to reduce costs have been decidedly mixed. Some studies have found little to no impact of HIT adoption on reducing costs and even found slight increases in average costs. Of the studies that have been previously discussed, Agha (2014) found no impact of the ability of

CDSS or electronic medical records to reduce costs or readmission and found an increase in billed charges.[12] In a similar fashion McCullough et al. (2013) find no impact in terms of reducing length of stay.

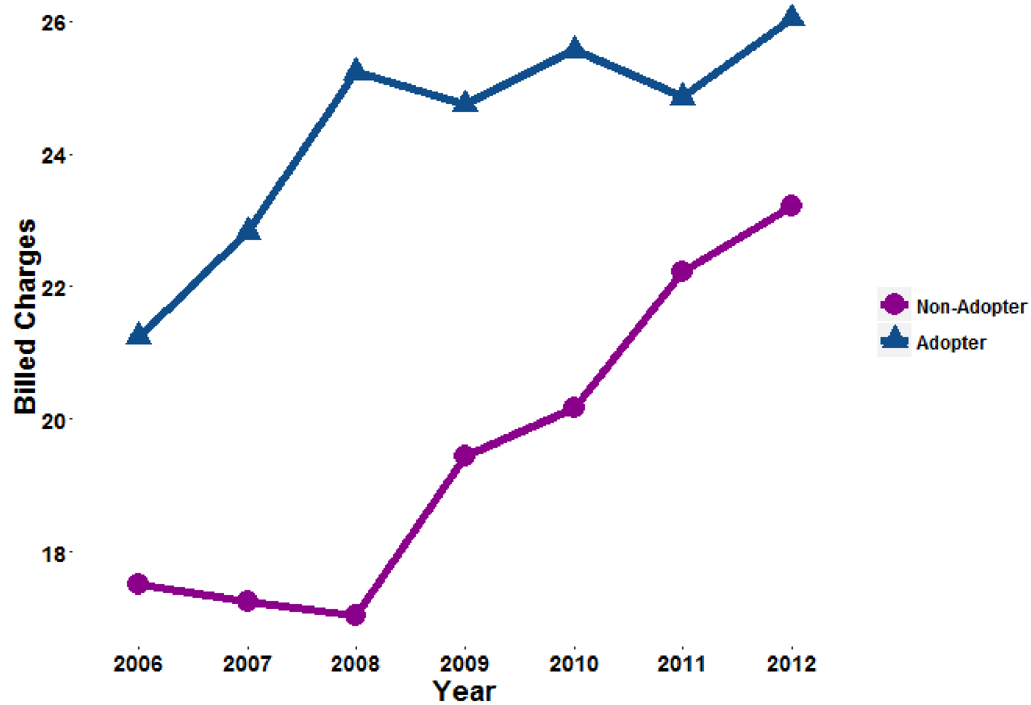
The somewhat disconcerting finding that HIT adoption is associated with an no decrease in costs and a potential increase in costs is puzzling and worrisome to say the least. However, Dranove et al. (2014) comment on this puzzling finding by drawing on information technology productivity literature. They find that in general, HIT adoption did not lead to lower average costs and in some cases led to higher average costs. However, they also highlight the importance of complementary environmental factors in the market in which the hospital resides. Important differences across hospitals (ex: location, available IT staff in local area, etc.) can have substantial impacts on costs, with hospitals in urban areas realizing lower costs three years after adoption. Hospitals in IT intense locations realized a 6.9% decrease in costs for basic HIT sytems and 7.3% for more advanced systems.

4.3 Extensions

When undertaking any analysis on costs or cost-side outcomes, such as length of stay or readmissions, it is imperative that great care be taken in the measurement of the outcome and the econometric specification employed. This is of paramount importance as there are many ways in which one may define "costs" and some metrics are more valid than others. For example, it is not sufficient to simply examine the ability of HIT to reduce billed charges for a hospital stay. Since what is billed is not necessarily reimbursed by the insurer or what the service cost the hospital, billed charges may be very misleading. For example, HIT components such as EMR, may make it easier to bill for additional services with the push of a button. However, this does not necessarily reflect was was reimbursed or reflect the cost on the patient. If one were to only examine this measure, they would almost certainly find a positive relationship between billed charges and HIT adoption, as Figure 4.1 shows.

A more appropriate metric may be one examining hospital operating costs such as Dranove et al. (2014). Or one may take a societal perspective and analyze the reduction of costs on consumers. The dataset utilized in this dissertation is fortunate enough to contain the billed charges which can be combined with hospital specific cost-to-charge ratios. These ratios will allow for the conversion of hospital specific charges to the cost of the procedure to

[12]It should be noted that billed charges may not accurately reflect the reimbursed cost that is negotiated by the insurer or the true cost of the treatment to the hospital.



Note: Dollars are in $100,000.

Figure 4.1: Billed Charges by HIT Adoption

the hospital and is common in the literature (Kazley et al., June 2014).

In addition to paying careful attention to how costs are defined, the underling theoretical modeling framework for outcomes linked to costs, such as readmissions and length of stay, need to be carefully specified. These outcomes are different from outcomes such as mortality because they are explicitly linked to costs or to policies that impose costs on hospitals. For example effective for discharges on October 1st, 2012 or later, the Affordable Care Act allows CMS to penalize hospitals with excessive readmissions. These penalties can be substantial, in the fourth year of the program 2,232 hospitals were penalized a collective $420 million. Also, the longer the length of stay for the patient, the more cost to the hospital so there is a financial incentive to discharge the patient as soon as clinically appropriate. Thus viewing these outcomes in a standard production framework, for example including exactly the same covariates as one would for mortality, is likely inappropriate.

Lastly, a consistent point of emphasis of this dissertation has been that

HIT adoption is endogenous. The fact that adopters are fundamentally different from non-adopters in terms of quality, size, and resources will certainly have cost implications. Higher quality facilities may have access to better IT staff and better training, shortening the learning curve and realizing HIT cost savings more quickly than lower-quality hospitals. Thus strategies that hinge upon the identification assumption that HIT adoption is unlikely to be correlated with unobservables seem inappropriate. Steps taken, such as those presented in this dissertation, are necessary for correcting the confounding nature of endogeneity and should continue to be presented.

5 Conclusions

This dissertation focused on the adoption of health information technology systems and the impacts of these systems on hospital inpatient outcomes. A novel dataset, which is not limited by payer type, was constructed and applied to both hospital level and patient level analyses. Another significant focus was the notion that HIT adoption is correlated with unobservable hospital level characteristics, and is thus endogenous. To combat this issue of endogeneity, the notion of using the HITECH Act of 2009 as a natural instrument was introduced.

The most important finding, is that I have found evidence to support that HIT systems are beginning to have significant impacts on improving patient outcomes. Compelling evidence was found at the hospital-level that HIT systems are reducing severity-adjusted mortality among Medicare patients. Additionally, for all three conditions examined (COPD, CHF, and pneumonia), HIT adoption significantly reduced the likelihood of dying regardless of the modeling strategy that was utilized. This points to the fact that HIT adoption is leading to improved outcomes, but that the impacts of HIT are different for different groups (Medicare, comorbidities, etc.).

References

Leila Agha. The effects of health information technology on costs and quality of medical care. *Journal of Health Economics*, 34:19–30, 2014.

Joshua D Angrist. Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business & Economic Statistics*, 19(1):2–16, 2001.

Manuel Arellano and Stehpen Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–297, 1991.

Kenneth Arrow. Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):941–973, 1963.

American Hospital Association. The road to meaningful use: What it takes to implement electronic health record systems in hospitals. Trendwatch: The American Hospital Association, April 2010.

Christopher F. Baum and Mark E. Schaffer. Enhanced routines for instrumental variables/gneralized method of moments estimation and testing. *The Stata Journal*, 7(4):465–506, 2007.

Christophe Bontemps and Cline Nauges. The impact of perceptions in averting-decision models: An application of the special regressor method to drinking water choices. *American Journal of Agricultural Economics*, 98 (1):297–313, 2015.

Melinda Beeukes Buntin and David Cutler. The two million dollar solution: Saving money by modernizing the health care system. *Center for American Progress*, 2009.

Akshay S. Desai and Lynne W. Stevenson. Rehospitalization for heart failure predict or prevent? *Circulation*, 126(4):501–506, 2012.

David Dranove, Forman Christopher, Avi Goldfarb, and Shane Greenstein. The trillion dollar conundrum: Complementarities and health information technology. *American Economic Journal: Economic Policy*, 6(4):239–270, August 2014.

David Dranove, Craig Garthwaite, Bingyang Li, and Chrisopher Ody. Investment subsidies and the adoption of electrionic medical records in hospitals. *Journal of Health Economics*, No w20553(44):309–319, 2015.

William H. Greene. Marginal effects in the bivariate probit model. NYU Working Paper No. EC-96-11, 1996.

DeFrances CJ Hall MJ, Levant S. Hospitalization for congestive heart failure: United states, 2000 to 2010. *Age*, 65(23):29, 2012.

Richard Hillestad, James Bigelow, Anthony Bower, Federico Girosi, Robin Meili, Richard Scoville, and Roger Taylor. Can electronic medical records transform health care? potential health benefits, savings, and costs. *Health Affairs*, 24(5):1103–1117, 2005.

Paul Hougland, Jonathan Nebeker, Steve Pickard, Mark Van Tuinen, Carol Masheter, Susan Elder, Scott Williams, and Wu Xu. Using icd-9-cm codes in hospital claims data to detect adverse events in patient safety surveillance. *Approaches in Patient Safety: New Directions and Alternative Approaches*, 1, 2008.

Ray GT et al. Huang SS, Johnson KM. Healthcare utilization and cost of pneumococcal disease in the united states. *Vaccine*, 28(18):3398–3412, 2011.

Anish Jah, Carherine DesRoches, Eric Campbell, Karen Donelan, Sowmya Rao, Timothy Ferris, Alexandra Shields, Sara Rosenbaum, and David Blumenthal. Use of electonic medical records in u.s. hospitals. *The New England Jounral of Medicine*, 360(16):1628–1638, 2009.

Gail A. Jensen and Michael A. Morrisey. The role of physicians in hospital production. *The Review of Economics and Statistics*, 68(3):432–442, 1986.

Abby Swanson Kazley, Annie N. Simpson, Kit N. Simpson, and Ron Teufel. Association of electronic health records with cost savings in a national sample. *The American Journal of Managed Care*, 20(6):e183–90, June 2014.

Arthur Lewbel, Yingying Dong, and Thomas Tao Yang. Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics*, 45(3):809–829, 2012.

J. S. McCullough, Michelle Casey, Ira Moscovice, and Shailendra Prasad. The effect of health information technology on quality in u.s. hospitals. *Health*, 29(4):647–654, 2010.

Jeffery S. McCullough, Stephen T. Parente, and Robert Town. Health information technology and patient outcomes: The role of organizational and information complementarities. *National Bureau of Economic Research*, No w18684, 2013.

Amalia R. Miller and Catherine E. Tucker. Can health care information technology save babies? *Journal of Political Economy*, 119(2):289–324, 2011.

2012 NCHS. Health, united states, 2012: With special feature on emergency care., 2013.

Barack Obama. Remarks of president-elect barack obama radio address on the economy, 2008.

OECD. Health at a glance 2011: Oecd indicators. OECD Publishing, 2012.

Michael Lokshin Richard C. Chiburis, Jishnu Dasb. A practical comparison of the bivariate probit and linear iv estimators. *Economics Letters*, 117(3): 762–766, 2012.

Jan Sidorov. It ain't necessarily so: The electronic health record and the unlikely prospect of reducing health care costs. *Health Affairs*, 25(4):1079–1085, 2006.

Douglas Staiger and James H. Stock. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586, 1997.

James H. Stock and Motohiro Yogo. *Identification and Inference for Economic Models: Essays in Honor of Thomas Rothenberg.*, chapter Testing for weak instruments in linear IV regression, pages 80–108. Cambridge University Press, 2005.

Elixhauser A. Pfuntner A. Au D.H. Wier, L.M. Overview of hospitalizations among patients with copd, 2008. statistical brief 106. Agency for Healthcare Research and Quality, February 2011.

WorldBank. Health expenditure. The World Bank Group, 2013. URL `http://data.worldbank.org/topic/health`.

Arnold Zellner, Jan Kmenta, and J. Dreze. Specification and estimation of cobb-douglas production models. *Econometrica*, 34(3):784–795, 1966.