ESSAYS ON THE ADOPTION OF ELECTRONIC MEDICAL RECORDS (EMR) BY U.S. HOSPITALS

by

Jianjing Lin

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Jianjing Lin, entitled Essays on the Adoption of Electronic Medical Records (EMR) by U.S. Hospitals
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 27 May 2015

Gautam Gowrisankaran

Date: 27 May 2015

Keith A. Joiner

Date: 27 May 2015

Mo Xiao

Date: 27 May 2015

Mauricio J. Varela

Date: 27 May 2015

Madhu Viswanathan

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.
I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 27 May 2015

Dissertation Director: Gautam Gowrisankaran

STATEMENT BY AUTHOR

SIGNED: Jianjing Lin

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

A $35 billion program was passed by the federal government to promote the adoption of Electronic Medical Records (EMR). However, billions of incentive payments were flowing out without clear evidence of effective implementation. The dissertation studies the adoption decision of EMR by U.S. hospitals and the consequence of the application of this technology. The first chapter tries to evaluate choosing the locally market-leading vendor by standalone hospitals. I construct a dynamic oligopoly model and apply the methodology developed by Aguirregabiria and Mira (2007) to recover the model primitives with a nationwide sample of U.S. hospitals. The primary finding is that, on average, the per-period profit from choosing the locally market-leading vendor is increased by almost 51% as opposed to that from using any other technology. However, the impact moderates as compared with the sunk cost of implementation. From the counterfactual analysis I find if hospitals were incentivized to choose the locally market-leading vendor, it would help improve the market coordination substantially.

The second chapter seeks to understand the incentive of hospital systems in choosing Health IT vendors: using the most-adopted product for coordination or otherwise to differentiate from the local market. I develop a simple discrete-choice model to evaluate the effect of each factor. Using a nationwide sample of affiliated hospitals from 2006 to 2010, I find that on average the system-dominating vendor has much greater advantage over the vendor leading the local market. After addressing the potential endogeneity issue, the impact from choosing the market-leading vendor is even negative. It may imply large systems are likely to create information silos, demonstrating lower propensity for external information exchange.

The last chapter examines the impact of adopting EMR on Medicare billing, particularly to understand how the application of Health IT affects hospitals' response to a recent payment reform. Using a nationwide sample of U.S. hospital and claims data, we find, in general, there is no significant difference in billing between hospitals with and without Health IT. However, hospitals behaved quite differently in documenting medical/surgical diagnoses before and after the reform.

CHAPTER 1

1.1 Introduction

This paper seeks to explore profit complementarities in adopting Electronic Medical Records (EMR) by stand-alone hospitals in the U.S. In particular, the study focuses on profit complementarities in choosing the market-leading vendor. EMR allow health care providers to store, retrieve, and exchange health information using computers instead of paper records. The advancement of this technology holds the promises to improve the efficiency of the health care system. In 2009, Congress passed legislation devoting $35 billion to promote the adoption of EMR, with the ultimate goal to allow seamless exchange of patient information. Almost four years and $12 billion later, data sharing, however, has proven very difficult to achieve. Recently, the government agency—Centers for Medicare and Medicaid Services (CMS)—announced its intention to change the timeline and revise the rulemaking for the last stage in the program, in order to accelerate information sharing. The purpose of my analysis is to provide some valuable inputs on the benefits of market coordination.

The definition of profit complementarities vary by context. In this case, it describes a situation in which a hospital can benefit from choosing a vendor that is also chosen by its neighbors. EMR systems are not interoperable. The system from one vendor cannot communicate with that from another. Hospitals may be more likely to extract profit complementarities if they use products from the same supplier. This paper looks specifically into profit complementarities from the local market-leading technology. There could be various sources. A vendor well-established in the local market may have developed a better relationship with local providers and payers. When hospitals process and submit claims, it will bring in great cost efficiency if its platform is compatible with the system of the payer. The leading supplier may also be able to provide sufficient expertise in the implementation of similar technologies. All of such could be translated into a cost advantage when the hospital purchases from the largest vendor.

A hospital choosing the most popular product can benefit from profit complementarities, but interoperability is not necessarily beneficial to a hospital. As indicated in an article by Kellermann and Jones (2013), health care providers have little incentive to "acquire and develop interoperable HIT systems" but would rather "lock in" their patients to "enforce brand loyalty." These two competing forces make it hard to theoretically determine the sign of using the market-leading technology. The objective of this paper is to examine which force dominates and to evaluate the effect of the market-leading technology on profits. By understanding this, regulators will be better positioned to expand the use of health IT and help hospitals make sound choices.

In 2009, the Health Information Technology and Economic and Clinical Health (HITE-CH) Act, a $35 billion program of grants and incentive payments, was passed as part of the American Recovery and Reinvestment Act (ARRA). The goal of this program was to establish a Nationwide Health Information Network (NwHIN) where patient information can be exchanged freely across diverse entities. Nearly four years after the enactment of the HITECH act, $12.7 billion has been paid out, but seamless transfer of health information is still out of reach. In the current incentive program, a hospital will be subsidized as long as it meets the specified requirements of meaningful use. The program doesn't set up any standard for interoperability nor impose any restriction on the choice of vendors. Thus, on one hand, hospitals are free to choose any vendor in the market. But on the other hand, since the products from different suppliers cannot talk with each other, the resulting information silos go against the original objective of the program and make the establishment of the NwHIN even more difficult. The government agency recently delayed the rulemaking for the last stage of the program and sought input on potential policy to accelerate information exchange across providers. At this moment when the policy makers reconsider the strategies needed to ensure interoperability, information about the value of the market-leading technology becomes important. If profit complementarities dominate the countervailing competition effect, promoting such a technology can not only bring in cost efficiency but is also helpful to improve coordination at the regional level. If the competition effect devaluates this technology, the government needs to take a closer look into the incentive behind hospitals' choices of vendors.

This paper contributes to studies on the adoption decision of EMR, most of which involve network externalities (Miller and Tucker (2009), Lee, McCullough and Town (2012)). The variable of interest, in all such studies, is the number of adopting hospitals nearby, to assess the extent to which the technology has permeated into the network. However, since EMR is not compatible, the presence of two adopting hospitals does not necessarily imply that both belong to the same network unless they are using the system from the same vendor. This paper addresses this issue by using the information of the identity of EMR vendors to define a network. In addition, it is the first study applying a dynamic framework to investigate the adoption choice of EMR vendors. In order to recover the value of using the market-leading technology, I develop a dynamic oligopoly model in the tradition of Maskin and Tirole (1988) and Erison and Pakes (1995). In this model, hospitals simultaneously make adoption decisions and the market evolves as hospitals adopt or switch to a new system in response to variation in the economic environment. The choice of vendors is a dynamic decision due to the following reasons. Large hospitals may have incentives to adopt early so that their choices become predominant in the market. They can benefit later on as more and more neighbor hospitals follow their choices. Small hospitals may hold up the current adoption decision, waiting to see which technology is optimal. Missing the dynamic element will omit all of these incentives. This is the first study making use of the information of the identity of EMR vendors to study the dynamics in hospitals' choice of vendor.

Using a nationwide sample of U.S. hospitals from 2006 to 2010, my estimation method follows the approach developed by Aguirregabiria and Mira (2007). I find that hospitals extract positive profits in selecting the market-leading vendor. On average, a hospital's profit from adopting EMR increases by almost 51% if it chooses the locally market-leading vendor. However, the effect moderates as compared with the substantial amount of the sunk cost. The gain is asymmetric between large and small hospitals. Profits generated from using the market-leading technology is higher for large hospitals, due to greater profit complementarities and less competitive pressure for these hospitals. I also find that both large and small hospitals have to bear a significant amount of switching cost when switching HIT vendors but large hospitals spend less in switching. A potential explanation is that large hospitals are capable of internalizing the cost of consulting and IT system management when switching. It dramatically reduces the expense to set up a different system since external consultancy and project management constitute the largest cost contributor.

With all the structural estimates, I am able to carry out the counterfactual analysis. I consider two types of subsidy programs. One is the subsidy for any type of EMR adoption, which attempts to mimic the current incentive program where the standard on interoperability is almost blank. The other one provides subsidies toward hospitals only if they choose the locally market-leading vendor. The outcome variable of interest is the rate of coordination in adopting EMR. The results suggest that encouraging the adoption of the market-leading technology could improve the market coordination substantially. I also compare the effect from the two programs in a new market, where the adoption rate is close to zero, and in a mature market, where almost all hospitals have adopted EMR but with almost zero communication. I find the targeted subsidy towards the most popular technology is much more effective in a new market. It is worthwhile to point out that promoting such a technology is just one way to set up the standard of interoperability. The key message is that the outcome is likely to have been better if the requirement on interoperability was explicitly incorporated at the earlier stage of the incentive program.

The rest of the paper proceeds as follows: Section 2 presents literature related to this topic. Section 3 provides basic information and institution background about EMR. Section 4 describes the datasets applied in this study. Section 5 provides reduced-form evidence on the value of using the market-leading technology. Section 6 presents the structure model characterizing the adoption decision of hospitals. Section 7 describes the estimation strategy. Section 8 shows the estimation results. Section 9 runs counterfactual experiments and discusses the potential policy implications. The last section concludes and points out the future directions.

1.2 Relation to Literature

Empirical studies examining the adoption decision of EMR usually involves network externalities. Miller and Tucker (2009) studied the relationship between privacy protection policy and technology diffusion. By comparing the states with and without the policy, their results suggested that privacy regulation inhibited the adoption of EMR by suppressing the network externalities. Another paper by Lee, McCullough and Town (2012) focused on the impact of Health Information Technology on hospital productivity and found little evidence of the network effect. Wang (2012) tried to disentangle the network externalities from the countervailing competitive effect by separately examining different adoption levels of EMR. She found that the basic level adoption yields a positive network effect while the advanced EMR application suggests a competitive effect. As indicated earlier, I define networks at the vendor level, unlike these studies.

A study by Dranove et al. (2012), from a different perspective, looked into the relationship between the hospital operational cost and EMR adoption. Their results indicate that hospitals benefit from EMR adoption when the necessary complements are in place; hospitals in a less favorable location undergo an increase in costs even after several years of installation. The supply condition of local complementary assets can help hospitals realize more profits from the adoption of EMR and hence they must be accounted for in the adoption decision. This paper emphasizes the role of the market-leading technology because one potential benefit from choosing such a product is that the vendor is more likely to supply sufficient complementary resources to its clients. This paper is also complementary to the empirical literature on network externalities. Tucker (2008) identified the network externalities from individual adoption of a video-messaging technology in an investment bank. Gowrisankaran and Stavins (2004) examined the extent of network externalities for automated clearing house (ACH). A follow-up study by Ackerberg and Gowrisankaran (2006) constructed an equilibrium model and structurally estimated the magnitude and sources. This paper incorporates the dynamic structure into a model in which a network is defined by using the technology from the same vendor.

A dynamic oligopoly model is constructed to characterize hospitals' adoption decision. There is a growing literature on estimating the dynamic models (Rust (1987); Hotz and Miller (1993); Hotz, Miller, Sanders and Smith (1994); Aguirregabiria and Mira (AM) (2003, 2007); Bajari, Benkard and Levin (2007); Pakes, Ostrovsky and Berry (POB) (2007); Arcidiacono and Miller (2011)). The seminal paper by Rust introduced the Nested Fixed Point (NXFP) algorithm for single-agent dynamic programming (DP) problems. It is a full solution method in the sense that the DP problem is solved for every trial value of the parameters. Moreover, under the assumption of the model, it gives the MLE that is asymptotically efficient. This method can be extended to estimate the dynamic oligopoly problem by assuming all the shocks are purely private information (Aguirregabiria and Mira (2007)). However, the limitation of the NXFP algorithm is the computational burden due to repeated full solution to the DP problem. Hotz and Miller (1993) observed the existence of the inverse mapping between the choice-specific probability and the difference in the choice-specific value functions. They proposed the Conditional Choice Probability (CCP) estimation method in which it is unnecessary to solve the DP problem even once to get the structural estimates. The key idea is to substitute the future values with future actions that can be nonparametrically estimated from the data. Aguirregabiria and Mira (2002) showed the asymptotic efficiency of the CCP estimator and suggested a recursive CCP algorithm to correct the possible inconsistency from the one-step CCP estimator. Arcidiacono and Miller (2011) extended the CCP framework from the model with a terminal state to a much wider set of dynamic problems.

Similar to Hotz ahd Miller's (1993) idea, Bajari, Benkard and Levin (2007) (BBL) provided a method for estimating the dynamic game models which also circumvents the needs to solve the Markov-perfect Nash Equilibrium (MPNE). The estimator proceeds in two steps. The first step is to estimate the policy function and the law of motion for state variables. In the second step, the structural estimates are recovered by imposing the optimality conditions for equilibrium. Ryan (2012) applied BBL to evaluate the welfare costs of the environmental regulation in a model where firms make the decision of investment, entry and exit. Collard-Wexler (2013) used a similar method in his paper to assess the role of demand shocks in the ready-mix concrete industry. The estimation strategy in this paper is based on a two-step framework, following closely the approach developed by Aguirregabiria and Mira (2007) but also similar to that in Bajari, Benkard and Levin (2007) and Pakes, Ostrovsky and Berry (2007). Hospitals are assumed to have correct belief about the environment and competitors' behavior so the policy function can be estimated from the equilibrium that is actually played in the data. With the estimated policy function, I apply forward simulation to get the value function.

The parameters are recovered by picking the values that are most probable to produce the observed behavior.

1.3 Industry Background

EMRs were invented in 1970s, but the acceptance to this technology had been very slow until recent years. In 2009, the American Recovery and Reinvestment Act (ARRA) has provided $35 billion to promote Health Information Technology (HIT), in particular to encourage the adoption EMR. It is the first substantial commitment of federal resources to support the adoption of EMR and creates a strong push in the diffusion of HIT. As the cornerstone of the Affordable Care Act in improving quality and lowering cost, EMR serves functions that paper record cannot deliver. According to the Healthcare Information and Management Systems Society (HIMSS), a solid EMRs foundation should include the following key components: Clinical Data Repository (CDR), Clinical Decision Support Capabilities (CDS), and Computerized Physician/Provider Order Entry (CPOE). CDR is essentially a centralized database that collects, stores, accesses and reports health information. It is the backbone of the entire system. CDS assists clinicians in decision-making tasks namely determining the diagnosis or setting treatment plans. CPOE is a more advanced type of electronic prescribing. It can link to the adverse drug event (ADE) system to avoid potential medication errors.

EMR has evolved from the early days being a silo system—in which the digital records from each ancillary department was isolated—towards nowadays an integrated architecture allowing sharing of data across departments, also known as the enterprise EMR system. The implementation cost of an EMR system varies tremendously depending on the sophistication of the system built, amount of data conversion, level of customization, one-on-one assistance during training and on-going use and etc. According to a study conducted by the Congress Budget

Office, the average implementation cost for a 250-bed hospital (about the mean size in my sample) ranges from \$3-\$16 million and the ongoing cost for subsequent upgrade and maintenance is approximately 20%-30% of the initial contract value per year, i.e., up to \$5 million annually afterwards. The rollout cost would even rocket to hundreds of million dollars for large hospitals. For example, in 2011 the medical center at the University of California, San Francisco spent \$150 million to have the EMR system in place. Such a large upfront payment involves a tangible part such as licence purchase, hardware investment, workflow consulting, project management and staff training. The last three components—external consultancy, project management and human capital investment—constitute the largest contributor in the upfront cost. The intangible cost mainly comes from the productivity loss during initial implementation. There are various reasons for hospitals willing to spend millions of dollars on this expensive technology. Besides the strong push from the federal incentive program, the usage of EMR enables hospitals to engage in better documentation, lower the administrative cost, and streamline and automate their revenue practices. Digitizing medical records also help hospitals get adapted to the reform in the payment system as well as the new features of the Accountable Care Organization. Last but not the least, a qualified EMR system can bring in more efficiency and improve the quality of health care.

However, the evidence about the effect of EMR adoption has been mixed. McCullough et al. (2010) connected health care quality to the use of CPOE and discovered substantive improvement from using the technology. Miller and Tucker (2011) provided a careful analysis of the impact on neonatal outcomes from the adoptions of EMR and found that a 10% increase in basic EMR adoption would reduce neonatal mortality rates by 16 deaths per 100,000 live births. Agha's job market paper (2011) investigated the impact of HIT on the quality and intensity of care delivered to Medicare patients but detected no significant improvement after the implementation. A more recent study by Li (2014) placed emphasis on the effect of EMR adoption on medical coding and billing practice. She used the multi-state inpatient discharge data to examine the relationship and found that the share of patients coded to a higher-pay DRGs increases significantly after EMR adoption.

The choice of EMR vendors relies on various factors such as the upfront and ongoing costs, the vendor-specific functionalities, individual hospital characteristics, payer impacts and local factors. In particular, the goal of this study focuses on profit complementarities from adopting the market-leading technology. The vendor with the highest local market share is defined as market-leading. In U.S., hospitals are divided into two categories according to the affiliation status: stand-alone hospitals that are independent organizations and affiliated hospitals that belong to a larger hospital chain. The analysis of this paper only concentrates on the sample of stand-alone hospitals while the adoption choice of affiliated hospitals is assumed to be exogenous to the local market.

I examine these hospitals for several reasons. For hospital systems, most organizational decisions are made by the managing party, who usually faces the tradeoff between localization and consolidation, especially when the locally-leading vendor is not the same as the choice of the parent system. Also, due to the heterogeneity in hospitals systems, it is difficult to characterize their decision process with a relatively simple model. By focusing on stand-alone hospitals, I can examine a relatively simple adoption decision but still identify the impact of profit complementarities. The data provides somewhat evidence in supporting this argument. Conditional on first-time adoption, only 19 % of affiliated hospitals chose the market-leading vendor while the rest followed the choice of the parent system even when it was not most widely-adopted in the local market.

1.4 Data

The data is constructed by pooling information from various sources. The first primary dataset comes from Healthcare Information and Management Systems Society (HIMSS) Analytics Database, which is the most comprehensive national source of hospital information technology (HIT) adoption data. The database covers the majority of U.S. hospitals, and includes market share and purchasing plan data for over 90 software applications and technologies. It is an annual survey recording the time and the choice of a hospital's adoption decision. More specifically, the dataset contains the information about the year adoption, the component deployed, adoption status, and the identity of the product supplier, which enables a more realistic network definition.

There is no consensus on how to define adoption of EMR for a hospital. Jha et al. (2009) used a very comprehensive definition. From a list of 32 potential functionalities of an inpatient electronic health record, they asked an expert panel to define the functionalities that constitute a basic and comprehensive electronic system respectively. Miller and Tucker (2009) measured EMR adoption by whether a hospital is installing or has installed the enterprise EMR system. In my paper, a hospital is defined to adopt EMR if CDR is live and operational in the hospital. The implementation of CDR is the prerequisite for other applications. It implies the hospital's willingness to enter the market and it is often the case that other typical and common applications such as CDS and CPOE will be put in place soon after the installation of CDR. This paper tries to understand the factors that will affect hospitals' choice of vendors. The adoption of CDR may uncover some information about hospitals' incentive. The second row of Table 2.5 and 2.6 report the nationwide adoption rate derived from the sample. In 2006 49% of the hospitals deployed EMR and the number went up to 84% in 2009. The adoption rate for stand-alone hospitals is slightly lower than that in general. The fraction of using the market-leading technology by stand-alone hospitals is more than 60% in both years.

I complement the HIT data with the American Hospital Association (AHA) Annual Survey, using the Medicare provider number and geographic information to perform the linkage. The AHA data includes a rich set of hospital-specific features such as number of beds, system affiliation, profit status, indicator of academic medical center, percentage of Medicare and Medicaid discharge and etc. Table 2.5 and 2.6 provides a summary statistics for the main variables. In 2006, about 45% of hospitals are stand-alone hospitals and this ratio fell by 2% in 2009. Also, the distribution of profit status and bed size remained almost the same during the sample period.

In this paper, a market is equivalent to a health service area (HSA), a measure developed by Makuc et al. (1991). An HSA is one or more counties that are relatively self-contained with respect to the provision of routine hospital care. The location of each hospital can be directly linked to the corresponding HSA. There are around 921 HSAs in the sample, covering more than 95% of HSA in US. The final dataset contains 4,560 hospitals between the year 2006 and 2009. Another thing to note is that the market for EMR is fairly concentrated. Although there are more 2,000 certified EMR vendors, most of the products are supplied by a few large companies. Table 2.4 lists the top 11 vendors that account for about 92% of the national market share in 2006. All the other vendors are categorized into another group called "others." Combining with the major vendors, these 12 options forms a choice set available to all hospitals in the model.

1.5 Reduced-form Evidence

This section provides some reduced-form evidence on how the hospital evaluates the market-leading technology. The two competing forces makes the sign of the value uncertain. To address this question, I run a conditional logit regression by regressing the choice of vendors on a set of product characteristics and hospital features. Whether the vendor is market-leading in the local area becomes a product characteristic entering into the value function. If the gain outruns the potential loss of business, the hospital will expect positive returns in choosing the market-leading technology and hence the estimated coefficient will be positive. If instead the concern about losing patients is greater, the coefficient is expected to be negative. The estimation is based on the sample of first-time adoption to provide a cleaner setting. Table 1.4 shows the results of the conditional logit regression. The upper panel reports the coefficients for the product characteristics: whether the vendor is market-leading and its interaction with the dummy for a large hospital. The latter aims to capture the extra gain/loss for large hospitals. A hospital is categorized as large if its number of beds is more than the mean size. The results suggests hospitals benefit from using the market-leading technology, implying profit complementarities exceed the loss from competition.

The lower panel presents the coefficients for the hospital characteristics interacting with vendor dummies. Bed size influences the choice of vendors as the variable cost (savings) could vary by product. In particular, the coefficients for Vendor CPSI and Healthland are negatively significant, indicating lack of cost savings per unit of bed. This is consistent with the design of these two products as both of them mainly target at rural and critical access hospitals most of which are small hospitals. Hospitals with different profit status may behave systematically differently in the choice of vendors. All of the coefficients for not-for-profit hospitals are positive, implying that such hospitals are more likely to adopt EMR. Teaching hospitals are responsible for clinical training and education for new generation of physicians in addition to delivering medical services, and thus they may have special preference in some particular vendors due to the various difference in functionality. The results suggest some vendors are much more popular among teaching hospitals. There could be other variables also playing some role in the choice of vendors, but the coefficients for the variables of interest—the product characteristics—seem quite robust across different specifications.

A potential concern about endogeneity may arise. There may exist some unobserved market characteristics affecting the formation of the leading technology and the choice of vendors at the same time. I use an outside market proxy as a cleaner measure. Consider a stand-alone hospital (called "A") in Chicago, which has, a competitor belonging to a hospital system for which most of the members locate in Phoenix. Instead of using the market structure in Chicago, I derive the product characteristic—whether a particular vendor is market-leading—from the market structure in Phoenix. The market condition in Phoenix is a clean measure in the sense that it is much less likely to be related to the market in Chicago, but it may have some association with the choice of Hospital A. The adoption decision of the hospital system may put some weight on the market in Phoenix, which will have impacts on the choice of its member in Chicago and further on Hospital A's decision. The idea here basically uses excluded variables from one system to identify another, similar to the strategy applied in the paper by Gowrisankaran and Stavins (2004). In order to apply this proxy, the sample is further restricted to the markets with both newly-adopting stand-alone hospitals and affiliated hospitals belonging to a multi-region system. This reduces the sample by one third. Table 1.5 reports the results in this specification. The market-leading dummy loses its significance, but its interaction with the large hospital dummy is positively significant. Large hospitals expect some gains from choosing the market-leading vendor.

1.6 Model

I develop a simple oligopoly model built on the theoretical framework constructed by Maskin and Tirole (1988) and Erison and Pakes (1995). There are M regional markets, each of which has N_m stand-alone hospitals $\forall m = 1, 2, ..., M$. Each market is fully described by an N_m state vector $s = (s_i, s_{-i})$ where s_i and s_{-i} are Hospital i's and its rivals' adoption status respectively, with $s_i = 0$ corresponding to no EMR in place and $s_i = j(j \in \mathcal{J} = \{1, 2, ..., J\})$ for using the product from Vendor j. $\mathcal{J} = \{1, 2, ..., J\}$ is a choice set which contains all EMR vendors available to hospitals. Since all the vendors are serving the national market, the choice set is fixed for every decision maker. Each hospital is assumed to capture a fixed portion of consumer surplus so oligopoly competition in medical care will not be explicitly considered in the model. For simplicity, I further assume there is no entry and exit in the market. Time is discrete and infinite. Each decision period is one year.

In each period, the sequence of events unfolds as follows: At the beginning of every period, each hospital receives a vector of private draws $\varepsilon = \{\varepsilon^0, \varepsilon^1, \varepsilon^2, ..., \varepsilon^J\}$ from some distribution. Conditional on a commonly observed vector of state variables s and their private shocks ε, all the hospitals simultaneously decide whether to adopt EMR and which vendor to choose. A hospital that has no EMR can either purchase from a choice set $\mathcal{J} = \{1, 2, ..., J\}$ or remains non-adopting. A hospital with some on-site system can either continue its own choice or switch to other vendors, but reversion is not allowed. A new purchase from Vendor j incurs a mean adoption cost ζ^j plus the unobserved component ε^j. ζ^j can be viewed as the net sunk cost of installing the software from Vendor j, which mainly includes the licence fee, upfront investment on hardware and human capital, resulting attrition and productivity loss. Switching from one vendor to another results in a switching cost η. It is modeled the same across vendors for identification and computational reasons. Switching did not occur very frequently during the sample period and inclusion of vendor-specific switching costs considerably increases the computational burden. A fixed amount of switching cost across all vendors may not be flexible enough to capture the vendor heterogeneity, but it is able to provides some idea about the mean expense for switching. Staying with the original choice results in the corresponding unobserved cost shock, such as ε^0 for non-adopting. Let ε terms be *i.i.d.* and have type I extreme value distributions. I assume it takes one period until the technology becomes operational within the organization. Thus, the action takes effect at the start of the following period and then the market evolves.

Consider Hospital i that has no EMR makes the purchase from Vendor j at period τ. At $t < \tau$, the per-period payoff is $\pi_{it}^{0}(s) = \varepsilon_{it}^{0}$. At the time of purchase $t = \tau$, the per-period payoff is

$$\pi_{it}^{j}(s) = -\zeta^{j} + \varepsilon_{it}^{j}. \tag{1.1}$$

When $t > \tau$, if Hospital i keeps using product j, the per-period payoff becomes

$$\pi_{it}^{j}(s) = \gamma_1 g(s,j) + \gamma_2^{j}\text{beds}_i + \gamma_3 g(s,j) \times 1\{i \text{ is a large hospital}\} + \varepsilon_{it}^{j} \tag{1.2}$$

where $g(s,j)$ is an indicator of whether Vendor j is market-leading at period t which can be calculated from the contemporary industry structure. γ_1 captures the profit complementarities if purchasing from the largest supplier and γ_2^{j} measures the vendor-specific cost savings per unit of bed. The term $\gamma_3 g(s,j) \times 1\{i \text{is a large hospital}\}$ captures the extra gain/loss for large hospitals if they choose the market-leading vendor. A positive γ_3 implies that the profit complementarities is increasing in the hospital's size. The hospital is viewed as large if the number of beds is greater than the mean size in the sample. Now I exposit Hospital i's decision problem in terms of the choice-specific value function (CSVF) $\delta^{j}(s)$. The choice-specific values represent the value of choosing each option absence of the unobservable component. Therefore, the CSVF for Hospital i to choose Vendor j is

$$\delta^{j}(s) = -\zeta^{j} + \beta \sum_{s'} P(s'|s, a(s) = j) EV(s') \tag{1.3}$$

and the CSVF to stay with no EMR is

$$\delta^{0}(s) = \beta \sum_{s'} P(s'|s, a(s) = 0) EV(s') \tag{1.4}$$

where β is the discount factor, ζ is the sunk cost of adoption, $EV(s')$ is the ex-ante future value function and $a(s)$ denotes the action chosen at state s. The transition probability, $P(\cdot|\cdot)$, depends on the firm's own behavior and equilibrium actions of its rivals. Therefore, the Bellman Equation for a hospital that has no EMR before is

$$V(s,\varepsilon) = \max_{j\in\{0\}\cup\mathcal{J}} \{\delta^j(s) + \varepsilon^j\}. \tag{1.5}$$

$V(s,\varepsilon)$ is the value function given the market state s and the private shocks ε faced by the hospital. By the same logic, I can write down the CSVF for Hospital i with product k chooses to continue its current choice

$$\begin{aligned} \delta^k(s) = \; & \gamma_1 g(s,k) + \gamma_2^k \text{beds} + \gamma_3 g(s,k) \times 1\{i \text{ is a large hospital}\} \\ & + \beta \sum_{s'} P(s'|s, a(s) = k) EV(s') \end{aligned} \tag{1.6}$$

while the CSVF for the hospital already with product k but deciding to switch to Vendor j is

$$\begin{aligned} \delta^j(s) = \; & -\zeta^j - \eta - \eta^{\text{big}} \times 1\{i \text{ is a large hospital}\} + \gamma_1 g(s,k) + \gamma_2^k \text{beds} \\ & + \gamma_3 g(s,k) \times 1\{i \text{ is a large hospital}\} + \beta \sum_{s'} P(s'|s, a(s) = j) EV(s') \end{aligned} \tag{1.7}$$

η measures the mean switching cost and the term $\eta^{\text{big}} \times 1\{i \text{is a large hospital}\}$ captures the extra gain/loss from switching for large hospitals. A positive η^{big} suggests that large hospitals have to bear a greater amount of switching cost. Similarly, the Bellman Equation is

$$V(s,\varepsilon) = \max_{j\in\mathcal{J}} \{\delta^j(s) + \varepsilon^j\}. \tag{1.8}$$

To summarize, the structural parameters involves the sunk cost ζ^j, the switching cost η and η^{big}, and the marginal values γ_1, γ_2^j and γ_3. The cost parameters measure the net present value while the γ's serve the flow pay off. Hospitals are assumed to play symmetric and Markovian strategies, i.e., the adoption choice only conditions on the current market state and the private shocks. Each hospital's adoption strategy is a mapping from state vectors and private shocks to the action:

$$\sigma : (s,\varepsilon) \to a. \tag{1.9}$$

Hospitals must weigh the benefits of using the new product against the adoption cost, the sum of the mean cost and the private draw. A purchase will not occur unless the sum of them is sufficiently low. The Markov-perfect Nash Equilibrium (MPNE) requires each hospital's adoption choice to be optimal given the strategy profiles of all rivals for all s, ε and all possible alternative choices $\tilde{\sigma}(s,\varepsilon)$. At least one MPNE exists according to the Brouwer fixed-point theorem. Pesendorfer and Schmidt-Dengler (2008) offered a nice proof of the existence. However, the uniqueness of the equilibrium is not guaranteed, which will be discussed more in the next session on empirical approach.

1.7 Empirical Strategy

The empirical strategy follows closely to the methodology developed by AM (2007), and it is also close to the approach in BBL (2007) and POB (2007). The first step is to estimate the equilibrium policy function. Assuming agents hold correct belief and play optimally, this step attempts to characterize hospitals' actions as a function of state variables. It avoids computing the equilibrium as the policy function is estimated from the equilibrium actually played in the data. The second step finds the parameters that rationalizes the observed policy function as the optimal choice given the underlying theoretical model. There are no guarantees that the equilibrium is unique. I impose the following assumption in order to group all the markets together in estimating the policy function.

Assumption 1: The same equilibrium is played in all markets.

This assumption is critical to obtain consistent estimates in the first step. Suppose there are two equilibria played in the data: $\sigma_1(s,\varepsilon)$ and $\sigma_2(s,\varepsilon)$. The estimated policy can be a convolution of both and therefore the imposition of the MPNE will generally not produce consistent estimates for the primitives. Under Assumption 1, I can group the markets together to recover the policy function.

Step One: Estimating the policy function

Hospitals make decisions about whether to adopt and which vendor to choose. The option actually picked should give the highest payoff. The probability for Hospital i to purchase from Vendor j is characterized using a logit regression:

$$P_i(a(s) = j) = \frac{\exp(x^j\alpha + z_i\lambda^j)}{\sum_{k=0}^{J}\exp(x^k\alpha + z_i\lambda^k)} \tag{1.10}$$

where x^j is a vector of product characteristics that vary by vendor and z_i represents a group of hospital-specific features. α's and λ's measure the marginal value of those variables. Note that the λ's vary across vendors. For example, if z_i includes the number of beds, the corresponding coefficient informs the marginal value per unit of bed from a particular vendor. In the model, the decision of affiliated hospitals is assumed to be exogenous to the local market but will also be observed by the stand-alone hospitals and thus enter the profit function. The evolution of these exogenous hospitals' choice is modeled analogously to that of the stand-alone hospitals. The x_j for stand-alone hospitals is a vector:

$$\begin{aligned} x_j = \ & [f(s,j),\ g(s,j),\ g(s,j) \times 1\{i \text{ is a large hospital}\}, \\ & l(j),\ l(j) \times 1\{i \text{ is a large hospital}\}] \end{aligned} \tag{1.11}$$

where $f(s,j)$ measures the market share of Vendor j; $g(s,j)$, defined as earlier, indicates whether Vendor j is the most popular in the local market; and $l(j)$ represents whether j is the same as the previous choice if Hospital i has already installed EMR. Both $f(\cdot)$ and $g(\cdot)$ are measures of the popularity for Vendor j. $l(\cdot)$ helps capture the inertia in adoption choice. Inclusion of $f(\cdot)$, $g(\cdot)$ and $g(\cdot) \times 1\{\text{i is a large hospital}\}$ helps to capture the transition of market states. As the number of possible states is more than the data points, I use those functions as approximation. Similarly, the x_j for affiliated hospitals includes

$$x_j = [r(j),\ l(j),\ l(j) \times 1\{i \text{ is a large hospital}\}] \tag{1.12}$$

where $r(j)$ indicates whether j is the major supplier for the entire hospital system. z_i could be different for both types of hospitals with different adoption status. According to the literature, important variables that affect the choice of vendors include environment factors like local competition levels, and hospital characteristics such as profit status, the number of beds, outpatient visits, inpatient admissions, full-time physicians, the percentage of Medicare and Medicaid discharges and whether the hospital is a teaching hospital. Which variables are included in the estimation depends on economic significance and model predictability. A market corresponds to a HSA and year combination. By assuming the errors to be *i.i.d.* across years and hospitals, I pool all the observations together for estimation.

Step Two: Recovering the structural parameters

The estimated policy function from the previous step describes how the state evolves over time and allows me to simulate the value of different choices which can be used to recover the model primitives. Starting with the actual state, I simulate the future market configurations for each action the hospital might take. States evolve according to the policy function estimated from the first stage. The simulated future paths will be long enough in order to approximate the infinite horizon problem (Here I use 100 periods). The value associated with a particular action will be the discounted values along the entire future path, i.e., the discounted sum of the expected per-period payoff from all future periods where the expected per-period payoff function is

$$\begin{aligned} E_{\varepsilon_{it}}\pi_{it}(s) = \ & [\gamma_1 g(s,k) + \gamma_2^k \text{beds} + \gamma_3 g(s,k) \times 1\{i \text{ is a large hospital}\}] \times 1\{\text{with } k \text{ in place}\} \\ & -\zeta^j \times 1\{\text{purchase from } j\} - \eta \times 1\{\text{switching from } k \text{ to } j\} \\ & -\eta^{big} \times 1\{i \text{ is a large hospital}\} \times 1\{\text{switching from } k \text{ to } j\} \end{aligned} \tag{1.13}$$

Another computational simplification in AM is that one doesn't have to do the forward simulation for every trial value of the parameters. Define

$$\begin{aligned} W(s_t;\sigma(s)) = \ & E_{\sigma(s)} \sum_{\tilde{t}=0}^{\infty} \beta^{\tilde{t}} [1\{\text{purchase from } j\},\ 1\{\text{switching from } k \text{ to } j\}, \\ & g(s_{t+\tilde{t}},k) \times 1\{\text{with } k \text{ in place}\},\ \text{beds} \times 1\{\text{with } k \text{ in place}\}] \end{aligned} \tag{1.14}$$

In my model, all the unknown parameters enter linearly into the payoff function. The value function is then

$$V(s_t;\sigma(s),\theta) = W(s_t;\sigma(s))\cdot[-\zeta^j,\ -\eta-\eta^{\text{big}}\times 1\{i\ \text{ is a large hospital}\},\ \gamma_1+\gamma_3\times 1\{i\text{ is a large hospital}\},\ \gamma_2^k]' \tag{1.15}$$

where θ denote the set of structural parameters to be estimated. By repeating such a simulation process for multiple times, the expected discounted value is the average across all the repetitions. Then I can write down the CSVF and hence the probability of each action. The structural parameters are estimated by maximizing the probability of observed actions.

1.8 Empirical Results

1.8.1 Policy function

The estimation of the policy function is essentially a multinomial logit regression and the main purpose is for forward simulation. Since hospitals that have no EMR have an extra option than those with an existing system, and since endogenous and exogenous hospitals have different policy functions governing the movement of their states, I run the conditional logit regression on four different subsets of the sample: stand-alone hospitals without EMR, stand-alone hospitals with EMR, affiliated hospitals without EMR and affiliated hospitals with EMR. As in Equation 1.11, the product characteristics x_j facing stand-alone hospitals contain the market share, whether a particular vendor is market-leading and its interaction with a large hospital dummy. The first two variables are different measures to reflect the popularity of each vendor and the last one helps capture the extra gain/loss for the large hospitals choosing the market-leading technology. The key role played by the policy function is to represent the transition across the states. The state space in this model can be enormous. The transition matrix for a market with 3 stand-alone hospitals (each with 13 options) is a 2137×2137 matrix, let alone the markets with more than 10 stand-alone hospitals. As the number of possible states is more than the data points, I use x_j as approximation. I also included into x_j other functions of the market states, such as whether it is the second or third leading technology and the interaction between all of them. The reported specification provides the best sample fit.

Table 1.6 displays the estimates of the policy function based on the sample of all stand-alone hospitals without EMR. The upper panel reports the coefficients for product characteristics and the lower presents the results for the hospitals characteristics interacting with vendor dummies. What hospitals features are included depends on economic significance and sample predictability. Vendors with higher market share seem to be more attractive to new adopters and using the market-leading technology is beneficial to large hospitals. The specification in this section is similar to that in the reduced-form estimation except that the sample applied here includes all stand-alone hospitals without EMR while the previous one is only a subset. The finding here is somewhat consistent with that in the previous specification with outside proxy. According to the lower panel, the probability of choosing a particular vendor increases with the bed size for most of the vendors except for CPSI and Healthland. This also confirms the previous finding that the products from these two supplier are better fit in smaller hospitals. All the vendor dummies are negatively significant, implying that new adopters are confronting substantial barriers in adopting the technology. Table 1.7 reports the results for the sample of stand-alone hospitals with EMR. In order to characterize the inertia of choices I add two more variables into the product characteristics: whether a particular vendor is chosen and its interaction with the large hospital dummy. The latter seeks to capture whether large hospitals behave differently than small ones. The last two columns in the upper panel show hospitals tend to be very "loyal" to the vendor they chose although large hospitals have a slightly higher chance to switch. A potential source of this "loyalty" may stem from a large amount of switching cost, which will be estimated in the second stage. The first three columns suggest that being the leading vendor plays a minor role in affecting the choice of hospitals with existing systems. The variables in the lower panel are somewhat different from that in the previous sample. What to include is again based on economic significance and model predictability.

The choices of affiliated hospitals evolves exogenously to the local market. However, their choices are part of the state variable and thus affect the adoption decision of stand-alone hospitals. I simulate the future choices of affiliated hospitals based on the "exogenous" product characteristics: an indicator of the hospital's system-wise dominant vendor and its previous choice. Specifically, the product characteristic for affiliated hospitals without EMR only involves an indicator of whether a particular vendor is dominant inside the hospital system. The upper panel in Table 1.8 provides the estimate of this variable and it is suggested member hospitals are more inclined to follow the choice of the parent system. The finding is robust for the set of affiliated hospitals with EMR, as shown in Table 1.9. This provides supportive evidence for the assumption that their decisions are exogenous to the local market. For the group of affiliated hospitals with EMR, I also include the two variables to capture the inertia and find that the previous choice plays a significant role in the adoption choice.

1.8.2 Structural Estimates

The policy function estimated from the first stage allows me to simulate the future states for all actions the hospital might take. The length of the future path is set to be 100 periods such that the discounted presented value of the last period is sufficiently small. I derive the value function by summing up all the future payoffs associated with each action and the probability is given by the assumed distribution of the unobserved shocks. Table 1.10 reports the estimates for the model primitives. The upper panel presents the value of using the market-leading technology, the switching cost as well as both interacting with the large hospital dummy. All of them are positively significant. Hospitals benefit from choosing the market-leading vendor and the gain is asymmetric between large and small hospitals. On average, profit generated from using the market-leading technology is 73.7% ($= 0.0157/0.0213 \times 100\%$) higher for large hospitals. There are two possible reasons: greater profit complementarities and less competitive pressure. A large hospital tends to interact more intensively with local payers and providers, and therefore using a compatible technology saves a lot of trouble. It is also more likely for them to get favorable pricing since they are often the preferred customers to vendors. When patients are able to switch between health care providers easily, a large hospital will expect inflow of patients given the advantage in technology and services. Both large and small hospitals bear a significant amount of switching cost but large hospitals spend 38.3% ($= 0.72/1.88 \times 100\%$) less in switching. External consultancy and system management are most expensive in implementation of EMR. As a result, large hospitals will set up their own department for IT support. If the hospital has to switch to a different vendor, the established IT team can be "recycled" to serve another system. However, this is not realistic for small hospitals, which instead has to keep hiring third-party consulting.

The lower panel in Table 1.10 presents the estimates of vendor-specific sunk cost and cost saving per 100 beds. The adoption of EMR incurs a considerable amount of sunk cost, regardless of which vendor to choose. The amount of sunk cost varies a lot by vendor. The lowest sunk cost is less than 40% ($= 2.71/6.80 \times 100\%$) of the highest one. If I roughly equate the median sunk cost (4.43) to the median implementation cost ($9.5 million according to the study by the Congress Budget Office) for a hospital with 250 beds, one unit of sunk cost represents about $2.2 ($= 9.5/4.43$) million. It is a very rough estimate since the sunk cost also includes the potential production loss which may not be accounted for in the pecuniary value. The third column in the lower panel in Table 1.10 lists the estimated cost savings for each vendor, most of which are positively significant except for the three vendors: CPSI, Healthland and HMS. It is consistent with the reduced-form evidence. For a hospital with 250 beds, adoption of EMR increases the per-period profit by 0.042 units from the cost-saving by bed size. If the EMR adopted is the market-leading technology, the hospital can further boost up its profit by 0.0213 units. Therefore, choosing the market-leading vendor can increase the per-period profit from adoption by 50.7% (= $0.0213/0.042 \times 100\%$) as opposed to other vendors. Although the market-leading technology brings in much higher payoff at each period, it is only 0.47% of the average sunk cost. Note that both the parameters for the market-leading technology and cost savings are based on one single period while the sunk cost measures the net discounted costs to implement the technology. In order to make the comparison more sensible, I adjust them into the net present values. Consistent with the discounted factor used in the estimation $\beta = 0.95$, a life-time payoff from using the market-leading technology is 0.43 (= $0.0215/(1-\beta)$) units, which accounts for 9.4% of the average sunk cost. The net gain from using the market-leading technology is moderate compared with the substantial cost barriers. The last two columns in the upper panel of Table 1.10 reveals the mean switching cost, which is almost 43% of the average sunk cost. This also explains why not everybody chooses the market-leading technology despite the potential benefit. At the moment of purchase, not all the hospitals have the perfect sense about which technology will be market-leading. If a hospital picked a choice that turns out to be suboptimal in the market, it would probably get stuck given the high switching cost. Note that all the analysis above are made for hospitals in general. Large hospitals would be probably in a more favorable position.

1.9 Counterfactual Analysis

Estimating a structural model allows me to simulate counterfactual experiments since I know the underlying primitives. My primary interest is to assess the impact on the adoption outcome across different policy regimes. To achieve this, I compute the MPNE of the theoretical model with the estimated parameters. The outcome variable is defined to be the rate of market coordination. It is the fraction of stand-alone hospitals that choose the market-leading technology which, additionally, is adopted by more than one hospitals. This extra requirement emphasizes the will of policy makers to coordinate hospitals' adoption choices. Consider a market with three hospitals, each installing a different system. Each hospital is using the market-leading technology, but no market coordination is occurring from the standpoint of the policy makers. It should be emphasized that this measure does not perfectly match the level of market coordination. Consider a different market with four hospitals. Two of them use the same product and the remaining share another one. In terms of this measure, the rate of coordination is 100% but in fact information cannot be exchanged freely in this market. The results should be interpreted in such a way that the emergence of this type of coordination reflect a certain level but not full degree of market coordination.

I should have solved out the full solution for each market in the data, but I only take the markets with three stand-alone hospitals into the analysis due to the computational constraints. It should be a reasonable measure as three active hospitals in a market is the average size in the data. Figure 1.1 compares the trend of adoption rate computed from the actual data with the in-sample prediction. The blue line shows the pattern generated from the data while the red line depicts the path predicted by the estimated model primitives. The model is doing a reasonable job for the in-sample prediction.

1.9.1 Too Many Choices?

This subsection explores a potential explanation for failure in market coordination. There are eleven major vendors plus hundreds of small ones available to hospitals, but each of them is not compatible with each other. In order to find out whether too many choices is one possible reason for poor coordination, I reduce the number of choices and compare the resulting rates of coordination. In the conducted experiment, I shut down all the vendors except for the six most popular ones and simulate the markets ten years forward. Figure 1.2 shows how the rate of market coordination evolves across time. The blue line describes the trend for the market with the full choice set while the red line for the experiment in which only six options are available. Both cases start with 18% of market coordination. Ten years later, it increases to 48% in the case of 6 choices while the markets with the original choice set have 34% coordination. The gap between two lines expands across time, implying having fewer options is helpful to improve the market coordination.

1.9.2 Subsidy in New Markets

The previous experiment provides some evidence that the number of choices has impacts on the level of coordination. I now perform counterfactual experiments under different policy regimes in new markets where most of the hospitals have no EMR. The first policy experiment involves the subsidy towards all adoption. As long as the hospital chooses to adopt, regardless of which vendor to pick, it will obtain a certain amount of subsidy. This unconditional subsidy program tries to mimic the element in the actual incentive program where no restriction is imposed on the choice of vendors. Given the fact that different products cannot communicate, the standard on interoperability is essentially blank under such a program. Another experiment considers a subsidy program in which hospitals have to choose the locally market-leading technology in order to get the reimbursement. It imposes an extra requirement, specially encouraging the adoption of the most popular technology in the local region. In each experiment, I derive the relationship between the amount of subsidy and the rate of market coordination across time. Due to the computation constraint, I restrict the available choices to be the six most popular vendors.

Figure 1.3 presents the relations under two policy regime. This is a 3-D graph with X axis being the amount of subsidy measured by the percentage of the median sunk cost, Y axis representing year and Z axis denoting the rate of market coordination. Both the left and right graphs represent the same figure from different perspectives. Particularly, the right one is the overlook of the graph. The green surface describes the evolution under the unconditional program and the red one presents that for the targeted subsidies. Whichever goes above represents greater market coordination. Taking a closer look at both graphs, I find the red surface slope goes steeper than the green one over time, implying that the targeted subsidies are even more effective over time. As time goes by, later adopters, who prefer to wait and see competitors' choices, will have better sense about which is the best technology in the market. Targeted subsidies towards the most popular technology make the choices for them even more obvious. Therefore, the market becomes increasingly integrated across time, hence widening the gap between the red and green surfaces.

However, the impact is not that obvious at the early stage. In the first two years, when the amount of subsidies is below 10 percent or between 38 to 58 percent of the median sunk cost, the unconditional subsidy program results in slightly better coordination. Intuitively, what distinguishes the targeted subsidies from the other lies in the fact that it forces hospitals to be more careful in the choice of vendors. Under such a program, early adopters tend to think twice about their choices so that the local leading vendor is likely to show up earlier. For later adopters, this extra requirement increases their option value of waiting and thus delays their action. The higher amount of subsidies is set, the more delay there will be. This is consistent with the finding that during the early years, the adoption rate under the unconditional subsidy program is increasingly higher in the amount of subsidies (shown in Figure 1.4).

When the amount of targeted subsidies is very low and during the first few years of the program, early adopters behave very much like they were in the program subsidizing any type of technology, and later adopters tend to delay their adoption just a little bit. Therefore, the targeted subsidies will not function very well and the outcome could be even worse than the other. As the amount of subsidies increases, early adopters are more careful in choosing the "right" vendor such that the local leading vendor "stands out" quickly. In the meantime, later adopters will not wait for too long and as a result, the market coordination in the presence of targeted subsidies catches up pretty soon and even outruns the other very early. When the amount of subsidies reaches a certain level where the sunk cost is no longer as a big concern as before, more and more hospitals are willing to adopt early in the case where any type of adoption gets subsidized. However, the generous subsidies will instead defer the adoption decision of even more hospitals in the targeted program. Therefore, early clustering is more likely to happen in the program offering unconditional subsidies when the amount of subsidies is rather high.

The graph on the right in Figure 1.3 shows the targeted subsidies to the locally leading technology dominates the other throughout the entire simulated period if the amount of subsidy is between 10% to 38% or more than 58% of the median sunk cost, which is between \$0.97 to \$3.90 million or more than \$5.85 million from the back-of-the-envelope calculation introduced earlier[1]. The red surface goes beyond the green one after 3 years as long as the amount of subsidies is more than 10% of the median sunk cost. In the actual incentive program, an eligible hospital with an average size (with 250 beds, 10,000 total discharge per year and 30% medicare discharge) can get \$1.2 million.[2] Assuming the amount from the model is a plausible estimate, the government is likely to have performed better with the same amount of expense. Promoting the market-leading technology is only one of the many measures to encourage a regional standard on interoperability. The comparison of these two experiments aims to illustrate the importance of

[1] For a hospital with 250 beds, one unit of sunk cost is equivalent to \$2.2 million.

[2] A simpler version of the official formula is that the incentive payment equals the product of the initial amount, percentage Medicare share discharge and transition factor. The initial amount is the sum of \$2 million plus \$200 per discharge for the 1,150th 23,000th discharge. The payment lasts for 5 years and I only use the subsidy in the first year so the transition factor is 1.

stressing the requirement about compatibility. In order to apply IT to health care effectively, the government should not only provide payments to purchase new technology, but can also take advantage of the profit complementarities to achieve interoperability. It is worthwhile to point out that I use the level of coordination as the outcome variable since the main policy concern is market integration. But hospitals' profits show similar results due to the fact that hospitals gain benefit from coordination. As shown in Figure 1.5, hospitals' profits are higher in the presence of targeted subsidies and the difference in profits becomes increasingly larger over time.

1.9.3 Subsidy in Mature Markets

This subsection evaluates the outcome from different types of subsidy in mature markets. Special attention is placed to markets with high adoption rate but almost no coordination. I design in the first experiment an unconditional subsidy towards all switching. Any hospital that switches can get a payment equal to the mean switching cost. The counterpart experiment imposes an additional requirement that hospitals will not get reimbursed unless they switch to the market-leading technology. Similarly, the latter experiment promotes the usage of the market-leading technology. I simulate ten years forward for both policy regimes and compare how the rates of market coordination change across time. Figure 1.6 plots the evolution for the case without subsidy, with unconditional subsidy and with subsidy towards special technology. All of them start with no market coordination. The rate of market coordination only increases to 24% after 10 years in the markets without any financial assistance from the government. In the first four years, both subsidy programs perform the same, but the one requiring extra condition outruns the other after then. The effect on mature markets is not as obvious as that in new markets. However, the government has to pay a higher price—the mean adoption cost is more than 40% of the mean sunk cost.

1.10 Conclusion and Future Work

This paper tries to understand the dynamics behind hospitals' adoption decision in the choice of EMR vendors. In particular, I focus on the sample of stand-alone hospitals which are considering whether to choose the technology leading the local market. On the one hand, hospitals gain profit complementarities from using the market-leading technology. However on the other hand, they are worried about losing patients when it is easier for them to switch between health care providers. The goal of this study is to estimate which force dominates and whether the policy makers can take advantage of this special property to improve market coordination.

I develop a dynamic oligopoly model to characterize hospitals' adoption decision. Whether a particular vendor is market-leading becomes a product characteristic goes into the profit function. Based on a national sample of U.S. hospitals, I estimate the structural parameters following the approach by Aguirregabiria and Mira (2007). It is found that hospitals expect positive returns from adopting the market-leading technology, i.e., profit complementarities exceeds the countervailing competitive effect. The gain is asymmetric between large and small hospitals. On average, choosing the market-leading vendor increases the per-period profit from adoption by almost 51% compared with other choices. However, the impact is moderate when it is compared with the amount of sunk cost in implementing the technology. Hospitals also have to bear a considerable amount of switching cost, about 43% of the average sunk cost, when they change vendors and small hospitals suffer even more. Consequently, it is important for the government to provide financial subsidy in assisting new purchase. However, the diffusion of HIT will be more effective if the government stresses the requirement of interoperability and one way is to incentivize the adoption of the market-leading technology. By recognizing the value of such a technology, the policy makers will play a significant role in helping hospitals make sound choices to improve market coordination.

The policy design in the current analysis is nondiscriminatory—a constant amount towards all hospitals. Considering the different magnitude in profit complementarities and switching costs between large and small hospitals, one can even go further by configuring some discriminatory payment mechanism to maximize the effect of the financial subsidy given the limited budget. I also consider modeling potential penalties, as percentage of total Medicare reimbursement, to hospitals which fail to demonstrate meaningful use of Health IT. It is believed that penalties in Medicare reimbursement could have even greater impacts on hospitals' incentive of adoption. Counterfactual experiments involving policy regimes with richer structure is part of the future plan for this project. I also consider including the supply side into the analysis. The current study models the hospital profit directly derived from adopting EMR without considering the vendors' behavior. The main reason is due to lack of data on the supply side. This part requires more extensive search in data and the literature.

Table 1.1: Summary statistics in the sample of 2006

Variable	Obs	Mean	Std. Dev.	Min	Max
EMR adoption for all hospitals	4560	0.49	0.50	0	1
EMR adoption for stand-alone hospitals	2060	0.41	0.49	0	1
% for stand-alone hospitals to use market-leading product	844	0.67	0.47	0	1
Number of hospitals in a HSA	4560	4.95	6.49	1	90
Number of stand-alone hospitals in a HSA	2060	2.24	2.66	0	33
Number of competitors in a local market	4560	3.95	6.49	0	89
Ratio of competitors adopting EMR	4560	0.37	0.37	0	1
Number of beds for stand-alone hospitals	2060	142.04	160.62	6	1558
For-profit stand-alone hospital	2060	0.04	0.20	0	1
Not-for-profit stand-alone hospital	2060	0.58	0.49	0	1
Academic medical center	2060	0.06	0.23	0	1
% of Medicare discharge for stand-alone	2060	0.50	0.15	0	0.99
% of Medicaid discharge for stand-alone	2060	0.16	0.11	0	0.84

Table 1.2: Summary statistics in the sample of 2009

Variable	Obs	Mean	Std. Dev.	Min	Max
EMR adoption for all hospitals	4560	0.84	0.37	0	1
EMR adoption for stand-alone hospitals	1969	0.77	0.42	0	1
% for stand-alone hospitals to use market-leading product	1517	0.64	0.48	0	1
Number of hospitals in a HSA	4560	4.95	6.49	1	90
Number of stand-alone hospitals in a HSA	1969	2.14	2.64	0	34
Number of competitors in a local market	4560	3.95	6.49	0	89
Ratio of competitors adopting EMR	4560	0.67	0.41	0	1
Number of beds for stand-alone hospitals	1969	140.04	166.38	4	2095
For-profit stand-alone hospital	1969	0.05	0.22	0	1
Not-for-profit stand-alone hospital	1969	0.56	0.50	0	1
Academic medical center	1969	0.06	0.23	0	1
% of Medicare discharge for stand-alone	1969	0.50	0.15	0	0.96
% of Medicaid discharge for stand-alone	1969	0.17	0.11	0	0.77

Table 1.3: Top 11 EMR vendors and their market share among in 2006

Vendor	Market share
Healthland	1.91%
EPIC	2%
Healthcare Mngt. Systems (HMS)	2.31%
GE Healthcare	3.33%
Eclipsys	3.87%
CPSI	5.69%
Self-developed (SD)	7.69%
Cerner	11.73%
Siemens	11.87%
McKessons	12.67%
Meditech	28.98%
sum	92.04%

Table 1.4: Reduced-form evidence—the existence of profit complementarities

market-leading	market-leading×big
0.6198***	0.0761
(0.1256)	(0.230)

	Cerner	CPSI	Healthland	Eclipsys	EPIC	GE
Bed size	0.0049	-0.0166**	-0.0190**	0.0120	0.0053	0.0113
	(0.0076)	(0.0077)	(0.0081)	(0.0075)	(0.0078)	(0.0079)
Not-for-profit	1.1884	0.6452	0.5785	2.1430*	1.9251*	2.3711*
	(0.8864)	(0.8258)	(0.8371)	(1.2343)	(1.0416)	(1.2731)
Teaching	15.1531***	-0.2143	-0.1803	-1.7297	16.1781***	13.7384***
	(1.6325)	(1.4762)	(1.5045)	(1.8380)	(1.5973)	(2.1080)
Constant	-0.1101	3.5495***	3.1470***	-3.0334***	-1.6027 **	-3.8790***
	(0.6959)	(0.6157)	(0.6293)	(1.1744)	(0.8097)	(1.0267)
	HMS	McKessons	Siemens	Meditec	Others	
Bed size	-0.0109	0.0065	0.0074	0.0003	0.0012	
	(0.0081)	(0.0075)	(0.0075)	(0.0074)	(0.0075)	
Not-for-profit	0.2397	0.9105	1.8690**	1.1484	1.1492	
	(0.8715)	(0.8379)	(0.8848)	(0.8149)	(0.8318)	
Teaching	0.0367	12.6544***	13.4320***	12.3908***	14.6913***	
	(1.5123)	(1.8003)	(1.6478)	(1.7802)	(1.5387)	
Constant	2.1529***	0.8840	-0.4512	2.3388***	1.5688 **	
	(0.6425)	(0.6347)	(0.7322)	(0.6182)	(0.6393)	
N=9768						

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 1.5: Reduced-form evidence—the existence of profit complementarities (using control function)

market-leading	market-leading×big
0.0846	2.6721***
(0.5007)	(0.8264)

	Cerner	CPSI	Healthland	Eclipsys	EPIC	GE
Constant	1.1231***	2.5824***	2.0422***	0.41	0.5092	-0.281
	(0.4058)	(0.3669)	(0.3753)	(0.4566)	(0.4492)	(0.5394)
	HMS	McKessons	Siemens	Meditec	Others	
Constant	1.4356***	1.8076***	1.3749***	2.5006***	2.0302***	
	(0.3933)	(0.3811)	(0.3939)	(0.379)	(0.3756)	
N=6780						

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 1.6: Estimates of the policy function (Stand-alone, w/o EMR)

market share	market-leading	market-leading×big
0.7122***	0.0044	0.4071**
(0.2453)	(0.1947)	(0.1904)

	SD	Cerner	CPSI	Healthland	Eclipsys	EPIC
Bed size	0.0029	0.0039***	-0.0081***	-0.0097***	0.0092***	0.0030
	(0.0043)	(0.0015)	(0.0017)	(0.0024)	(0.0010)	(0.0022)
Not-for-profit	-0.3980	0.6693*	0.3075*	0.2335	1.6830**	1.3250*
	(0.8363)	(0.4063)	(0.1748)	(0.2215)	(0.8505)	(0.7686)
Teaching	-13.8569***	0.9334	-13.3599***	-13.3339***	-14.5723***	2.1994**
	(1.6132)	(0.9720)	(0.3572)	(0.4449)	(0.7770)	(0.9278)
Constant	-5.6129***	-5.1636***	-2.3821***	-2.8037***	-7.6426***	-6.4284***
	(0.4253)	(0.3077)	(0.1196)	(0.1586)	(0.8595)	(0.4848)
	GE	HMS	McKessons	Siemens	Meditec	Others
Bed size	0.0086***	-0.0039	0.0059***	0.0062***	0.0018**	0.0023**
	(0.0021)	(0.0025)	(0.0008)	(0.0010)	(0.0007)	(0.0010)
Not-for-profit	1.6069	-0.1314	0.4959**	1.4263***	0.7558***	0.7334***
	(0.9951)	(0.3494)	(0.2528)	(0.3613)	(0.1527)	(0.2237)
Teaching	-0.6056	-13.3982***	-2.0211**	-1.3320	-2.0329*	0.0326
	(1.3673)	(0.5704)	(1.0385)	(0.8674)	(1.0846)	(0.7409)
Constant	-8.2930***	-3.6769***	-4.2640***	-5.4798***	-3.1045***	-3.8352***
	(0.6505)	(0.1962)	(0.1874)	(0.3508)	(0.1244)	(0.1829)
N=39624						

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 1.7: Estimates of the policy function (Stand-alone, w/ EMR)

market share	market-leading	market-leading×big	same_as_previous	same_as_previous×big
0.3027	0.0802	0.019	5.8852***	-0.4140*
(0.3872)	(0.2388)	(0.2773)	(0.182)	(0.2235)

	Cerner	CPSI	Healthland	Eclipsys	EPIC	GE
Bed size	0.0013	-0.0129***	-0.0196***	0.001	0.0016	0.0023*
	(0.0011)	(0.0022)	(0.0031)	(0.0016)	(0.0013)	(0.0014)
For-profit	-1.3010**	0.5389	-1.5226***	-1.4885***	-2.6347***	0.1332
	(0.6575)	(0.4868)	(0.462)	(0.5538)	(0.693)	(0.8156)
Teaching	0.3018	-10.5295***	-10.6777***	1.0249	0.8907	0.8781
	(0.8891)	(0.777)	(0.7898)	(0.9378)	(0.918)	(0.87)
Constant	0.8503**	2.7640***	3.0426***	0.3158	1.0952**	-1.7326***
	(0.4173)	(0.4343)	(0.5161)	(0.5835)	(0.4712)	(0.6666)
	HMS	McKessons	Siemens	Meditec	Others	
Bed size	-0.0136***	0.0003	0.0015	-0.0009	-0.0005	
	(0.0042)	(0.0014)	(0.0012)	(0.0012)	(0.0013)	
For-profit	-1.2161*	-2.2365***	-0.6236	-2.4632***	-0.0096	
	(0.686)	(0.7165)	(0.4669)	(0.9237)	(0.4624)	
Teaching	-10.1295***	0.2218	0.1521	-1.4998*	0.734	
	(0.9406)	(0.961)	(0.8238)	(0.8419)	(1.0517)	
Constant	1.9324***	1.0681**	0.0895	2.3956***	-0.388	
	(0.4937)	(0.508)	(0.4849)	(0.4163)	(0.4505)	
N=60324						

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 1.8: Estimates of the policy function (Affiliated, w/o EMR)

system dominating
2.0853***
(0.0884)

	Self-developed	Cerner	CPSI	Healthland	Eclipsys	EPIC
Bed size	0.0019***	0.0021***	-0.0069**	-0.0143***	0.0020*	0.0012*
	(0.0007)	(0.0006)	(0.003)	(0.0042)	(0.0012)	(0.0007)
Not-for-profit	-0.8441**	0.4582**	-0.5896*	-0.0908	1.6677**	3.2469***
	(0.3774)	(0.2226)	(0.3348)	(0.3563)	(0.7533)	(0.7101)
%Medicare	-1.2401*	-1.6060**	1.4855	0.0684	-2.1646	-1.5231*
	(0.6365)	(0.686)	(1.2944)	(1.2139)	(1.9756)	(0.8769)
Constant	-3.5953***	-3.0245***	-3.9492***	-2.9596***	-5.3726***	-5.5964***
	(0.429)	(0.4199)	(0.8509)	(0.809)	(1.2677)	(0.8697)
	GE	HMS	McKessons	Siemens	Meditec	Others
Bed size	0.0024***	-0.0026*	0.0020***	0.0004	0.0005	0.0014*
	(0.0009)	(0.0015)	(0.0005)	(0.0007)	(0.0006)	(0.0008)
Not-for-profit	2.2271***	-1.4520***	0.7611***	0.0887	1.2992***	0.7282**
	(0.7336)	(0.3466)	(0.2035)	(0.2075)	(0.1911)	(0.3312)
%Medicare	-0.8437	-2.3751***	-1.3543**	-3.9831***	-1.1476*	-1.5962
	(1.2326)	(0.5267)	(0.6567)	(0.5953)	(0.6338)	(0.9959)
Constant	-6.1285***	-1.4750***	-3.3206***	-1.5589***	-3.8481***	-3.8598***
	(1.1405)	(0.3537)	(0.3795)	(0.2993)	(0.4008)	(0.6114)
N=34944						

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 1.9: Estimates of the policy function (Affiliated, w/ EMR)

system dominating	same_as_previous	same_as_previous×big
1.3089***	4.8336***	-0.2557*
(0.0957)	(0.0987)	(0.1345)

	Cerner	CPSI	Healthland	Eclipsys	EPIC	GE
Bed size	0.0019***	-0.0044***	-0.0073**	0.0028***	0.0028***	0.0017**
	(0.0007)	(0.001)	(0.003)	(0.0007)	(0.0006)	(0.0007)
%Medicare	3.7011***	4.3433***	5.4168***	3.4056**	3.1445***	2.193
	(1.0458)	(1.0596)	(1.7236)	(1.5248)	(1.0454)	(1.4086)
%Medicaid	2.5690*	3.2605**	1.9818	1.1805	-2.0057	-0.3332
	(1.5042)	(1.4944)	(3.9046)	(2.3911)	(1.8089)	(1.7571)
Constant	-2.1377***	-2.3350***	-3.4792**	-2.6049***	-1.1816*	-1.8963**
	(0.6027)	(0.6464)	(1.4056)	(0.9931)	(0.6655)	(0.8419)
	HMS	McKessons	Siemens	Meditec	Others	
Bed size	-0.0049***	0.0011	0.0002	0.0001	0.0012	
	(0.0013)	(0.0008)	(0.0007)	(0.0007)	(0.001)	
%Medicare	3.9610***	3.8027***	1.7096	3.7860***	6.1945***	
	(1.4713)	(1.3155)	(1.1398)	(1.1509)	(1.5661)	
%Medicaid	4.4903**	4.3568**	2.0883	3.1581*	5.1797***	
	(1.7558)	(1.7312)	(1.6912)	(1.738)	(1.8224)	
Constant	-2.3549**	-2.8050***	-1.9669***	-2.2121***	-4.9706***	
	(0.9196)	(0.8024)	(0.7314)	(0.6859)	(0.9676)	
N=89724						

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

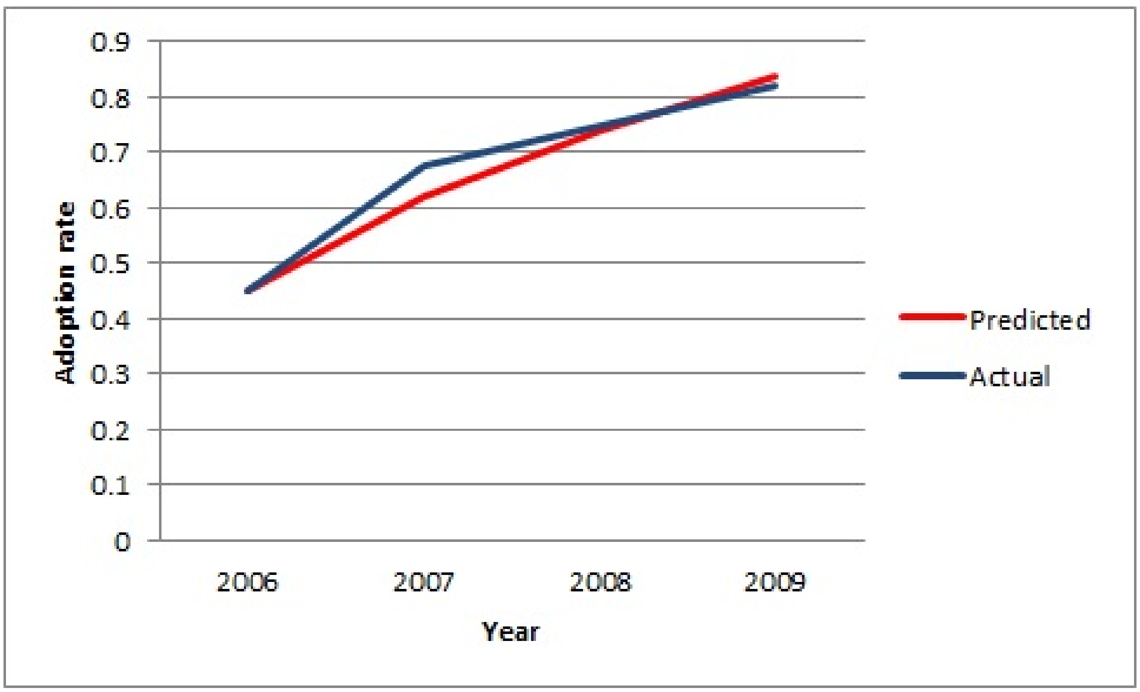
Table 1.10: Structural estimates for the dynamic model

market-leading	market-leading×big	switching cost	switching cost×big
0.0213***	0.0157*	1.8834***	-0.7199***
(0.005)	(0.0082)	(0.14)	(0.1738)

	Sunk cost	Cost saving per 100 beds
Self_developed	5.8169***	0.018***
	(0.2945)	(0.0041)
Cerner	5.2462***	0.0362***
	(0.2452)	(0.0047)
CPSI	2.7055***	-0.0158
	(0.1338)	(0.0103)
Healthland	3.0657***	-0.0303*
	(0.1724)	(0.016)
Eclipsys	5.6912***	0.0374***
	(0.3177)	(0.0042)
EPIC	5.5308***	0.0418***
	(0.2661)	(0.0039)
GE	6.8015***	0.0406***
	(0.475)	(0.0049)
HMS	4.0098***	-0.0152
	(0.2219)	(0.0161)
McKessons	4.1194***	0.0299***
	(0.1627)	(0.0034)
Siemens	4.7431***	0.0301***
	(0.1902)	(0.0036)
Meditec	3.0472***	0.0199***
	(0.13)	(0.0039)
Others	3.5761***	0.0101***
	(0.1267)	(0.0031)

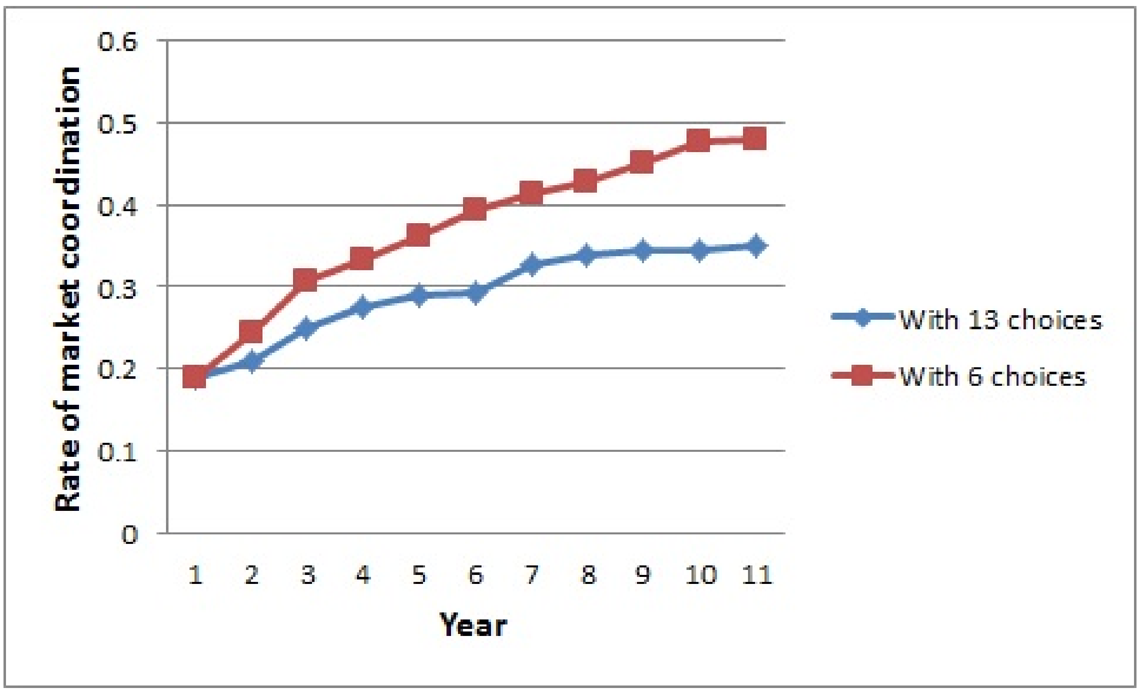
Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Figure 1.1: In-sample prediction of adoption rate



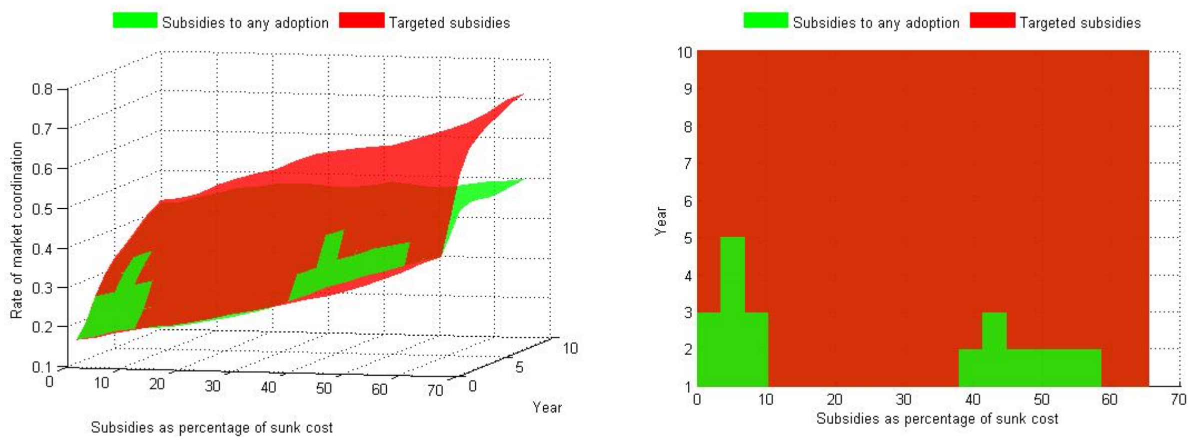
Note: Figure 1.1 compares the trend of adoption rate computed from the actual data with the in-sample prediction. The blue line shows the pattern generated from the data while the red line depicts the path predicted by the estimated model primitives. The model is doing a reasonable job for the in-sample prediction.

Figure 1.2: Market coordination between different choice sets



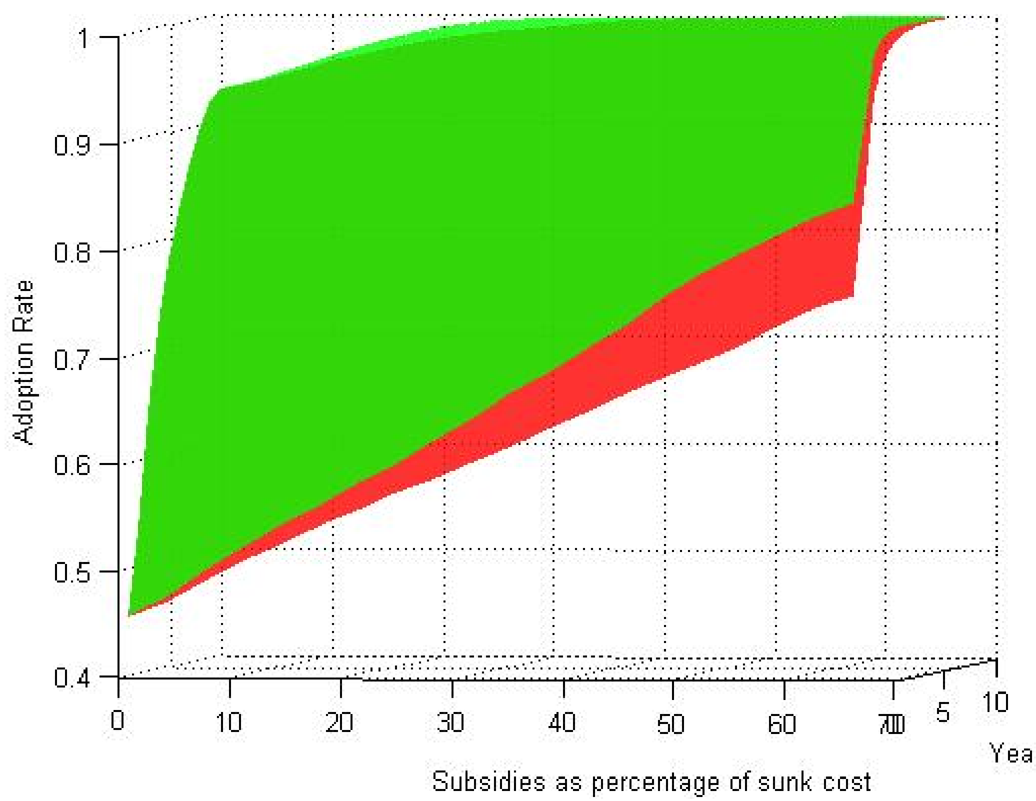
Note: Figure 1.2 shows how the rate of market coordination evolves across time. The blue line describes the trend for the market with the full choice set while the red line for the experiment in which only six options are available. Both cases start with 18% of market coordination. Ten years later, it increases to 48% in the case of 6 choices while the markets with the original choice set have 34% coordination. The gap between two lines expands across time, implying having fewer options is helpful to improve the market coordination.

Figure 1.3: Market coordination under different subsidy regime—new markets



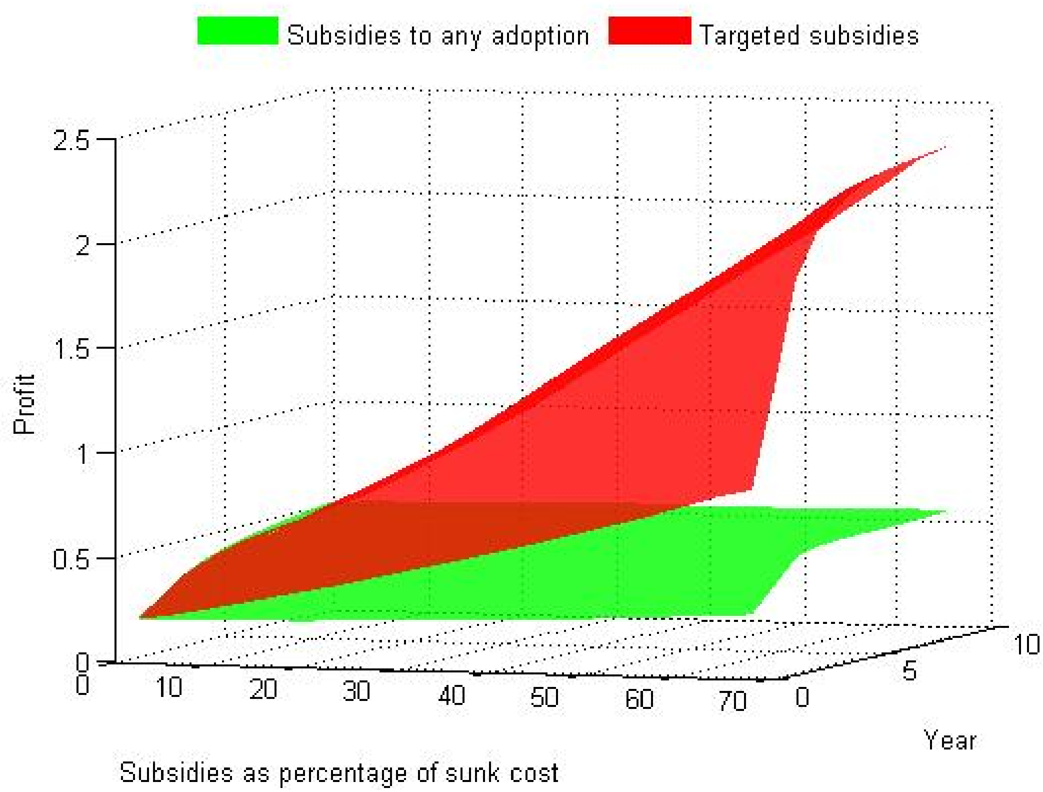
Note: Figure 1.3 presents the relations under two policy regime. This is a 3-D graph with X axis being the amount of subsidy measured by the percentage of the median sunk cost, Y axis representing year and Z axis denoting the rate of market coordination. Both the left and right graphs are the same figure from different perspectives. Particularly, the right one is the overlook of the graph. The green surface describes the evolution under the unconditional program and the red one presents that for the targeted subsidies. Whichever goes above represents greater market coordination.

Figure 1.4: Adoption rate under different subsidy regime—new markets



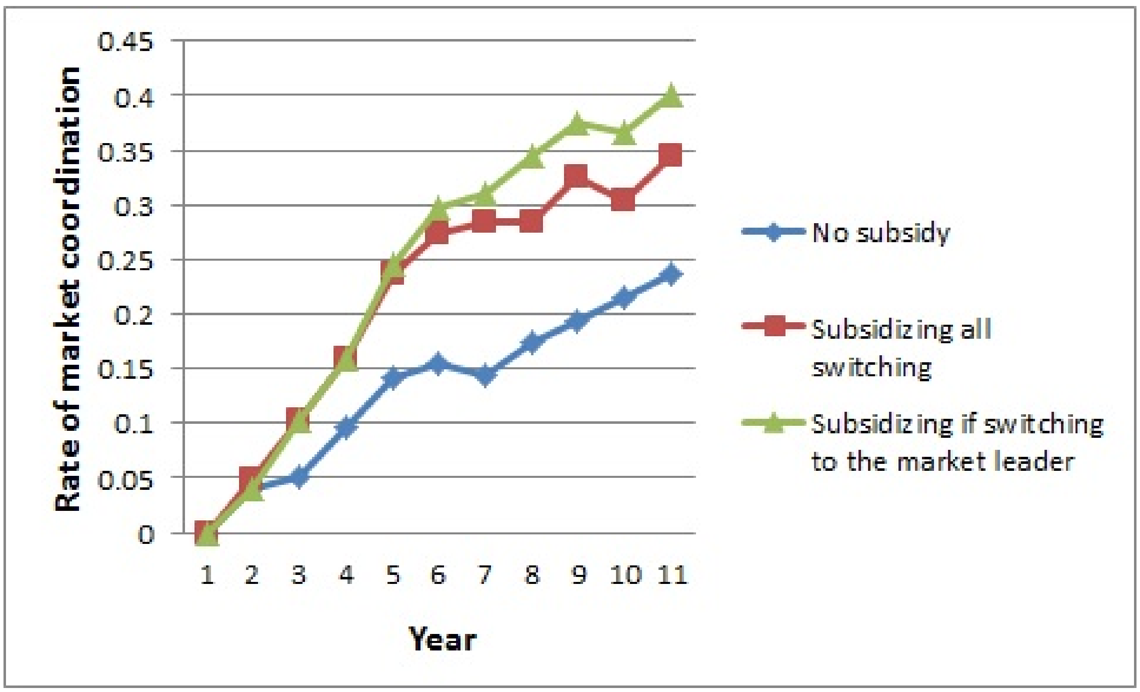
Note: Figure 1.4 presents the relations under two policy regime. This is a 3-D graph with X axis being the amount of subsidy measured by the percentage of the median sunk cost, Y axis representing year and Z axis denoting the adoption rate in the market. The green surface describes the evolution under the unconditional program and the red one presents that for the targeted subsidies. Whichever goes above represents higher adoption rate.

Figure 1.5: Profits under different subsidy regime—new markets



Note: Figure 1.5 presents the profits under two policy regime. This is a 3-D graph with X axis being the amount of subsidy measured by the percentage of the median sunk cost, Y axis representing year and Z axis denoting profits in the market. The green surface describes the evolution under the unconditional program and the red one presents that for the targeted subsidies. Whichever goes above represents higher profits.

Figure 1.6: Market coordination under different subsidy regime—mature markets



Note: Figure 1.6 plots the evolution for the case without subsidy, with unconditional subsidy and with subsidy towards special technology. All of them start with no market coordination. The rate of market coordination only increases to 24% after 10 years in the markets without any financial assistance from the government. In the first four years, both subsidy programs perform the same, but the one requiring extra condition outruns the other after then.

CHAPTER 2

2.1 Introduction

This paper explores the adoption decision of Electronic Medical Records (EMR) by hospital systems in U.S.. The decision of these affiliated hospitals could face such a tradeoff in choice, between the market-leading vendor, which has the highest local market share, and the system-dominant vendor, which is chosen by most member hospitals in the same system. Knowing how each factor impacts profits is important for policy makers to understand the incentive of hospital systems in adoption and thus to be better positioned in improving the coordination of adopting Health IT.

EMR allows healthcare providers to store, retrieve and exchange health information using computers instead of paper records. It serves functions that paper record cannot deliver such as improved efficiency through lower administrative cost, increased safety via accurate drug utilization and better health outcomes from disease prevention and chronic-disease management. Almost every industry is computerized for rapid data retrieval and information analysis. However, the speed of transmission to digitized records had been surprisingly slow in the field of medicine. Jha et al. (2008) indicated that other industrialized countries such as the United Kingdom, Netherlands, Australia and New Zealand have almost 100% adoption of ambulatory EMR by primary care physicians while the adoption rate in US only ranged from 10% to 30%. In 2009, the Health Information Technology and Economic and Clinical Health (HITE-CH) Act, a \$27 billion program of grants and incentive payments, was passed as part of the American Recovery and Reinvestment Act (ARRA). The goal of this program was to establish a Nationwide Health Information Network (NwHIN) where patient information can be exchanged freely across diverse entities. Nearly four years after the enactment of the HITECH act, \$12.7 billion has been paid out, but seamless transfer of health information is still out of reach. The purpose of my analysis is to provide some valuable inputs in understanding the dynamics behind the adoption decision by affiliated hospitals, which is important for future policy design in ensuring the occurrence of information exchange between health care providers.

This paper is the first economic analysis to make use of detailed information on the identity of the health IT vendor. The EMR system from one company is not compatible with another, but systems from the same vendor are able to communicate with relative ease. Recognizing the lack of interoperability, hospitals need to decide whether to adopt and further to make a choice among vendors, i.e., which network to join. In particular, the decision confronting a member hospital belonging to a hospital chain involves whether to follow the choice of the parent system or to choose the locally market-leading vendor, especially when both are different. This paper takes one step back by recognizing the fact of incompatibility and tries to understand which choice is more valuable from the perspective of a hospital: the market-leading vendor or the system-dominant vendor. Selecting the most widely-adopted vendor in the local market implies the hospital is joining the largest local network, which can possibly bring in lots of benefits. For instance, a vendor well-established in the local market may have developed a better relationship with local providers and payers. When the hospital processes and submits claims, it will bring in great cost efficiency if its platform is compatible with the system of the local payer. The leading supplier may also be able to provide sufficient expertise in the implementation of similar technologies. All of such could be translated into cost advantage when the hospital purchases from the largest local vendor.

However, EMR may be also regarded as a competitive technology. As indicated by Kellermann and Jones (2013) , "healthcare providers have had little incentive to acquire or develop interoperable health IT systems". When information can be seamlessly transferred, hospitals, on the one hand, may expect inflows of data from neighboring hospitals. However, on the other hand, hospitals are likely to fear the loss of patients like those with Preferred Provider Organization medical insurance as these insurance policies set less stringent rules for referral. Miller and Tucker (2014) studied the relation between health information exchange and the system size of hospitals. They showed that larger hospitals are more likely to exchange information internally but less willing to share information externally, reflecting a tendency to create "information silos". An information silo is a data system incapable of cross-talk with other systems. In fact, affiliated hospitals have the incentive to follow the choice of the home system to achieve economies of scale. Installing and operating an EMR system not only involves a substantial amount of fixed cost but also incurs enormous expenditure on technological assistance and system maintenance. According to a study conducted by the Congress Budget Office, the average implementation cost for a 250-bed hospital (about the mean size in my sample) ranges from \$3-\$16 million. The subsequent ongoing cost usually takes 20-30% of the initial investments per year. Using the product that has already been established in the parent system could save a great amount of cost. For instance, member hospitals could share the licensing fee, human-capital training expense, and the expenditure on external consultancy about operation and maintenance for IT systems with similar designs. The saving could be significant as the latter two constitute the largest cost contributor.

It is important to understand the incentive behind the adoption decision for hospital systems. More than 55% of hospitals are part of a hospital chain and they are of large size on average. Thus, the adoption choice of these hospitals is likely to influence the decision of other independent health care providers in the local area. Aware of the driving force of the decision for them is critical for policy design in facilitating health information exchange. If the market-leading product brings in greater value, hospitals tend to coordinate and internalize the potential network externalities. If instead hospitals are inclined to deviate from the local market, it reflects a tendency for hospitals to create "information silos". Given the stake of billion dollars on this technology, learning about the underlying reason of adoption is important for the government to achieve its goal. In this paper, I apply a simple discrete-choice model of adoption where hospitals make a choice between vendors. Whether the vendor is market-leading and whether it is system-dominating become two important characteristics that affect the profit from adoption. The result of the analysis suggests that the system-dominant vendor is much more preferable than the other. In some specifications when endogeneity has been addressed, the value of choosing the local market leader becomes even negative. It is shown that being consolidated with the parent system in adopting EMR is much more advantageous.

The rest of the paper proceeds as follows: Section 2 presents literature related to this topic. Section 3 provides basic information and institutional background about EMR. Section 4 describes the datasets applied in this study. Section 5 presents a simple model characterizing hospitals' choice of EMR vendors. Section 6 describes the estimation strategy. Section 7 shows the estimation results. The last section concludes and points out the future directions.

2.2 Relation to Literature

Empirical studies examining the adoption decision of EMR usually involve network externalities. Miller and Tucker (2009) studied the relationship between privacy protection policy and technology diffusion. By comparing the states with and without the policy, their results suggested that privacy regulation inhibited the adoption of EMR by suppressing the network externalities. Another paper by Lee, McCullough and Town (2012) focused on the impact of Health Information Technology on hospital productivity and found little evidence of the network effect. A more recent study is the job market paper by Wang (2012) who tried to disentangle the network externalities from the countervailing competitive effect by separately examining different adoption levels of EMR. It is found that the basic level adoption yields a positive network effect while the advanced EMR application suggests a competitive effect. In all such studies, the variable of interest is the number of adopting hospitals nearby, which is used to assess the extent to which the technology has permeated into the network. However, since the EMR systems from different vendors are not interoperable, the presence of two adopting hospitals does not necessarily imply both belong to the same network unless they are using the system from the same vendor. I use the information of the identity of EMR vendors to define a network. A hospital that purchased the software from a particular vendor also expects information exchange with the providers who share the same vendor.

A study by Dranove et al. (2012), from a different perspective, looked into the relationship between the hospital operational cost and EMR adoption. Their results indicate that hospitals benefit from EMR adoption when the necessary complements are in place; hospitals in a less favorable location undergo an increase in costs even after several years of installation. The supply condition of local complementary assets can help hospitals realize more profits from the adoption of EMR and hence they must be accounted for in the adoption decision. This paper emphasizes the role of the market-leading technology because one potential benefit from choosing such a product is that the vendor is more likely to supply sufficient complementary resources to its clients. This paper is also complementary to the empirical literature on network externalities. Tucker (2008) identified the network externalities from individual adoption of a video-messaging technology in an investment bank. Gowrisankaran and Stavins (2004) examined the extent of network externalities for automated clearing house (ACH). A follow-up study by Ackerberg and Gowrisankaran (2006) constructed an equilibrium model and structurally estimated the magnitude and sources. One identification strategy applied in this paper is similar to that by Gowrisankaran and Stavins (2004).

2.3 Industry Background

EMR were invented in 1970s, but the acceptance to this technology had been very slow until recent years. In 2009, the American Recovery and Reinvestment Act (ARRA) has provided $27 billion to promote Health Information Technology (HIT), in particular to encourage the adoption EMR. It is the first substantial commitment of federal resources to support the adoption of EMR and creates a strong push in the diffusion of HIT. As the cornerstone of the Affordable Care Act in improving quality and lowering cost, EMR serves functions that paper records cannot deliver. According to the Healthcare Information and Management Systems Society (HIMSS), a solid EMR foundation should include the following key components: Clinical Data Repository (CDR), Clinical Decision Support Capabilities (CDS), and Computerized Physician/Provider Order Entry (CPOE). CDR is essentially a centralized database that collects, stores, accesses and reports health information. It is the backbone of the entire system. CDS assists clinicians in decision-making tasks namely determining the diagnosis or setting treatment plans. CPOE is a more advanced type of electronic prescribing. It can link to the adverse drug event (ADE) system to avoid potential medication errors.

There is rich literature on the effect of the adoption of EMR in the health care sector. McCullough et al. (2010) connected healthcare quality to the use of CPOE and discovered substantive improvement from using the technology. Miller and Tucker (2011) provided a careful analysis of the impact on neonatal outcomes from the adoptions of EMRs and found that a 10% increase in basic EMR adoption would reduce neonatal mortality rates by 16 deaths per 100,000 live births. Agha (2011) investigated the impact of HIT on the quality and intensity of care delivered to Medicare patients but detected no significant improvement after the implementation. The conclusion on the external effect of EMR also reached no consensus.

This paper studies the adoption behavior of hospitals and focuses on two types of product characteristics: whether it is the most popular among local hospitals and whether it is the major Health IT provider of the hospital system. The vendor with the highest local market share is defined as market-leading and the vendor that is implemented by most members is called the system-dominating vendor. Note that both the local market leader and system-dominant vendor are defined without considering the hospital's own choice. Suppose there are 3 hospitals in a market: A, B and C, with A adopting product X, B adopting product Y and C without any IT system. For hospital A, the local market leader is the vendor producing Y. For hospital B, it is the one selling product X and for hospital C, both X and Y are the leading technologies. It is determined likewise on whether a particular vendor is dominating within a hospital system. Such a definition is helpful to reduce some endogeneity issues in estimation.

A hospital chain/system is defined by AHA "as either a multihospital or a diversified single hospital system. A multihospital system is two or more hospitals owned, leased, sponsored, or contract managed by a central organization. Single, freestanding hospitals may be categorized as a system by bringing into membership three or more, and at least 25 percent, of their owned or leased non-hospital preacute or postacute health care organizations"[1]. For hospital systems, most organizational decisions are made by the managing party, who usually faces the tradeoff between localization and unification, especially when the market-leading vendor is not the same as the system-dominant one. The distribution of their choices may reveal some information about how the decision maker evaluates each feature. Table 2.1, 2.2, and 2.3 present the distribution

[1]The definition comes from the Fast Facts Archive 2014 from the AHA Annual Survey. See http://www.aha.org/research/rc/stat-studies/fast-facts.shtml

of adoption choices for all the systems, the multi-region systems and the single-region systems respectively, conditional on the case that the set of system-dominant vendors is exclusive to that of market-leading vendors. More specifically, I select a group of hospitals for which the market-leading vendors are different from the system-dominant vendors. According to whether the chosen vendor is market-leading and whether it is system-dominant, its choice will fall into one of the following categories: choosing the local market leader, choosing the major vendor for the parent system or neither. In general, almost 70% of affiliated hospitals select the system-dominant vendor while less than 10% choose to adopt from the local market leader. In particular for systems locating in only one area, more than 84% of member hospitals choose to follow the parent system's choice but only about 4% choose the other. The results suggest that on average affiliated hospitals are more inclined to follow the choice of the parent system, when it is distinct from the most popular vendor. I use a simple t-test for the hypothesis that the percentage using the market-leading product is that same as the percentage following the home system's choice. The null hypothesis is rejected in all the cases. It may imply hospitals are more inclined to select a different product from the current market instead of coordinating with the local network, but one cannot conclude without more rigorous analysis as there are various factors affecting the adoption decision.

Another point to note is that the market of EMR is fairly concentrated. Although there are more than 2,000 certified EMR vendors, most of the products are supplied by a few large companies. Table 2.4 lists the top 11 vendors that account for about 92% of the national market share in 2006. All the other vendors are categorized into an additional group called "others". Combining with the major vendors, these 12 options form a choice set of EMR vendors available to all hospitals in the model.

2.4 Data

The data is constructed by pooling information from various sources. The first primary dataset comes from Healthcare Information and Management Systems Society (HIMSS) Analytics Database, which is the most comprehensive national source of hospital information technology (HIT) adoption data. The database covers the majority of U.S. hospitals and includes market share and purchasing plan data for over 90 software applications and technologies. It is an annual survey recording the time and the choice of a hospital's adoption decision. More specifically, the dataset contains the information about the identity of the product supplier, enabling a more realistic network definition.

There is no consensus on how to define adoption of EMR for a hospital. Jha et al. (2009) used a very comprehensive definition. From a list of 32 potential functionalities of an inpatient electronic health record, they asked an expert panel to define the functionalities that constitute a basic and comprehensive electronic system respectively. Miller and Tucker (2009) measured EMR adoption by whether a hospital is installing or has installed the enterprise EMR system. In my paper, a hospital is defined to adopt EMR if CDR is live and operational in the hospital. The implementation of CDR is the prerequisite for other applications. It implies the hospital's willingness to enter the market and it is often the case that other typical and common applications such as CDS and CPOE will be put in place soon after the installation of CDR. This paper tries to understand the factors that will affect hospitals' choice of vendors. The adoption of CDR may be able to uncover some information about hospitals' incentive. The first row of Table 2.5 and 2.6 report the nationwide adoption rate derived from the entire sample. In 2006 49% of the hospitals deployed EMR and the number went up to 84% in 2009. The adoption rate for affiliated hospitals was slightly higher, from 56% to 89% during the sample period. Among affiliated hospitals that have adopted EMR, a higher proportion chose the system-dominant vendors than those using the market-leading products.

I complement the HIT data with the American Hospital Association (AHA) Annual Survey, using the Medicare provider number and geographic information to perform the linkage. The AHA data includes a rich set of hospital-specific features such as number of beds, system affiliation, profit status and etc. Table 5 and 6 provide summary statistics for the main variables. In 2006, around 54.8% of hospitals are affiliated to a hospital chain and this ratio rose by about 2% in 2009. Also, the distribution of variables representing hospital characters like profit status and bed size remained almost the same during the sample period.

In this paper, a market is equivalent to a hospital service area (HSA), a measure developed by Makuc et al. (1991). An HSA is one or more counties that are relatively self-contained with respect to the provision of routine hospital care. The location of each hospital can be directly linked to the corresponding HSA. There are around 921 HSAs in the sample, covering more than 95% of HSA in US. The final dataset contains 4,560 hospitals between the year 2006 and 2010.

2.5 Empirical Model

In this section, I model hospitals' discrete choice decisions over EMR vendors. This paper only focuses on the adoption choice of affiliated hospitals, since the main question addressed in this paper is to understand how hospitals evaluate internal consolidation versus external coordination. The result of the analysis can provide some potentially useful inputs for policy makers to improve the coordination of adopting HIT. I start with this group of hospitals not only due to the fact that they happen to face such a trade off but also for the reason that they are on average of larger size and hence possibly more influential players in the regional market.

There are M regional markets, each of which has N_m affiliated hospitals $\forall m = 1, 2, ..., M$. These affiliated hospitals associate with K hospital systems and each contains N_k members $\forall k = 1, 2, ..., K$. In other words, $\sum_m N_m = \sum_k N_k$. To maximize the ex ante profit, every hospital must simultaneously decide whether to adopt EMR and if yes, the hospital also needs to choose a supplier j from a set of vendors $\mathcal{J} = \{1, 2, ..., J\}$ including the major vendors listed in Table 2.4 plus the additional group "others". Given that all the vendors are serving the national market, the choice set is fixed for every decision maker. The set $\mathcal{J}$ plus the outside option of not-adopting form a choice set available to all the hospitals. A hospital that has no EMR can either purchase from a choice set $\mathcal{J} = \{1, 2, ..., J\}$ or remains non-adopting. A hospital with some on-site system can either continue its current choice or switch to other vendors, but reversion is not allowed. Note that in reality the decision of each member hospital can either be formed by itself or be instructed by the managing party, but the analysis is conducted based on revealed preferences. In other words, the hospital can make the decision on its own or follow the instruction from the managing party. The essence of this model is to uncover how different factors are evaluated from the action actually taken by hospitals. Each hospital is assumed to capture a fixed portion of consumer surplus so oligopoly competition in medical care will not be explicitly considered in the model. For simplicity, I further assume there is no entry and exit in the market. Time is discrete and infinite.

More precisely, hospital i's profit from adopting vendor j at time t is given as follows:

$$\pi_{it}^j = X_{it}^j \alpha + Z_{it} \gamma^j + u_{it}^j \tag{2.1}$$

where X_{it}^j includes a set of variables capturing the vendor characteristics. Table 2.7 lists variables included in X_{it}^j. Whether vendor j is locally market-leading and whether it is the dominant EMR provider for the chain hospital i belongs to are the two key variables of interest. Inclusion of whether a particular choice is the same as the previous period aims to capture the inertia in adoption. The remaining variables are interactions with the hospital and system size. A hospital is defined to be large if its bed size exceeds the mean size in the sample. I use the number of HSAs in which a system is located and the number of its members as a proxy for the size of a hospital system. Particularly, the former provides a measure of how "spread out " the system is. I also use the total number of beds of a system but the results are not significant in all specifications. These interaction terms aim to pick up the extra impacts for large hospitals/hospital systems. There are other features that may affect the profit from adoption, such as price, interface design, functionality, availability of secure messaging, referral management feature and etc. I use vendor fixed effects to pick up these factors. However, if how these features influence profits is time-varying, one has to interpret the results with caution.

Consistent with the literature, Z_{it} contains important variables that affect the choice of vendors, such as environment factors like local competition levels, and hospital characteristics such as profit status, the number of beds, outpatient visits, inpatient admissions, full-time physicians, the percentage of Medicare and Medicaid discharges and whether the hospital is a teaching hospital. The profit shock, u_{it}^j, has mean zero and follows some *i.i.d.* distribution. Hospitals select a vendor in their choice set which generates the highest profits. The profit for non-adoption is normalized to zero. I can impose different types of distributions on the profit shock. Consider a hospital without Health IT, whose profit shocks have a Type I extreme value distribution. Let

$$\delta_{it}^j = X_{it}^j \alpha + Z_{it} \gamma^j \tag{2.2}$$

denote the "mean profit" for choosing vendor j at time t. Given such an assumption on the error distribution, the choice probability of adopting vendor j is

$$P_{it}^j = \frac{\exp \delta_{it}^j}{\sum_{k \in \{0\} \cup \mathcal{J}} \exp \delta_{it}^k} \tag{2.3}$$

for hospital i if it hasn't adopted any technology. By the same logic, I can write down the choice probability for hospital i if it has already adopted EMR.

2.6 Empirical Strategy

This section describes the empirical strategy to estimate how affiliated hospitals evaluate vendor characteristics: being market-leading, being system-dominant, as well as whether these impacts vary across different types of hospital systems. With different assumptions on the error distribution, I can either estimate the model with maximum likelihood or use the linear probability model (LPM). In this paper, I report the results from the LPM for several reasons. The foremost reason is to maintain consistency as I use the Two-Stage Least Squares (TSLS) to address the potential endogeneity issue in the later part of the analysis. Additionally, it has the advantage of easily-interpreted coefficients as the procedure involves no intermediate steps to calculate the marginal effects.

Each observation corresponds to a hospital and year combination. For each observation, I stack all the available choices together. Consider a hospital without any Health IT in 2006. The dependent variable for this observation is a vector with 13 entries, each of which equals one if the corresponding option is chosen and zero otherwise. The independent variables include both product and hospital characteristics. The variables of product characteristics have the same dimension in the sense that each entry describes the feature of a particular vendor. For instance, each entry of the indicator for the local market leader is equal to one if the corresponding vendor is most popular in the local market. I interact the hospital characteristics with vendor dummies, aiming to control for the varying effect of hospital features across vendors. Consequently for each observation, the data is stacked by combining all the available choices together.

I estimate this model with Ordinary Least Squares (OLS). Table 2.7 lists the variables capturing vendor characteristics. The basic idea of identification is to examine whether hospitals are more likely to cluster around the local market-leading vendor or the system-dominant Health IT provider. If I find clustering around the local market leader, I will interpret it as willingness of external coordination. If instead more hospitals are found to follow the parent system's choice, internal integration plays a more important role and it is likely to hurt market coordination when the system-wise choice is different than the most popular vendor in the local market.

However, a potential concern may arise about the endogeneity from the indicator of the market leader. There may exist some unobserved market characteristics, such as market-wise promotion or special preferences of local physicians, affecting the formation of the leading technology and the choice of vendors at the same time. For instance, a vendor based in this market may provide promotion to all local hospitals. The leading vendor becomes the most adopted simply because of the promotion rather than the benefit from coordination. To address this issue, the first set of instruments is created by averaging the market share and the indicator of market dominance for each vendor across the outside associated markets. Consider a market M_1 with three hospitals A, B and C. The endogenous variable for A is the vector with each entry denoting the market leadership for the corresponding vendor. Suppose B is affiliated to a hospital chain $Beta$, most of whose members locate in outside markets M_2 and M_3, i.e., the majority of members locate in M_2 and M_3 and the number in each market is the same. Both M_2 and M_3 are called the outside associated markets. Suppose C belongs to another hospital chain $Zeta$, whose majority members locate in M_1 and (a different market) M_4. Thus, one of the instruments for A is the average of the market leader indicators across M_2, M_3 and M_4. In addition, I use another instrument by averaging the market share of each vendor across the outside associated markets. This set of instruments are relevant in the sense that both the managing parties of *Beta* and *Zeta* may consider the market condition in these areas and thus affect the choice of B and C, which may have impacts on A's decision. It could be a clean measure as it is plausible that these outside markets have little relation with the unobservables in M_1. The idea here basically uses excluded variables from one system to identify another, similar to the strategy applied in the paper by Gowrisankaran and Stavins (2004). However, the application of the first set of instruments depends on the assumption that there are no spillovers across.

I use the second set of instruments to further rule out the inter-market dependence. Suppose a vendor offers special rates to all hospitals in M_1 and the promotion can be extended to the entire hospital chain. If the number of hospitals of *Beta* dominate M_2, this outside market is no longer exogenous. To address this issue, I look for single-region hospital systems locating in the outside associated markets. From Table 2.3, only about 4% of hospitals from single-region systems choose the local market leader when it is different from the system's major provider, implying their decision is much less affected by the local market situation. As a result, it seems to be a measure with less interdependency across markets but still relevant to the endogenous variable. Likewise for each of these "qualified" single-region hospital chains, I calculate the share of each vendor adopted by members and the indicator describing system dominance for each vendor. The second set of instruments are generated by averaging across these "qualified" hospital chains. Note that I focus on the endoneity issue at the local market level without considering that arising from the hospital chain. There could exist some unobservable factors, other than complementarities, that will affect the choice of vendors and the generation of a system-dominant vendor, such as special promotion from a vendor or specific strategy at the chain level. I do not try to separate them at this stage as such profit-enhancing behavior at the chain level is what I attempt to capture from this variable. Without separating these factors, the estimated coefficient reveals the general value of choosing the system-dominant vendor, which could result from complementarities between member hospitals and, more importantly, the profit-enhancing motive of the chain. This conveys important messages about hospital chains' incentive in choosing EMR vendors.

There has been a strand of literature on dealing with continuous endogenous variables in the discrete choice model. However, it may not be valid to apply these methods on my model since the endogenous variable in this model is discrete. According to Wooldridge (2002), those approaches depend on the assumption of continuous distributions on the error term, which has ruled out the application on any discrete endogenous variable. Therefore, the most straightforward approach to get to this issue is the LPM with TSLS. This may not be a terrible option as suggested by Angrist and Pischke (2009:198-204) and much empirical evidence. For robustness check, I also use ivprobit and control function to address the endogeneity issue and the results are quite consistent.[2] A market corresponds to a HSA and year combination. The profit shocks are assumed to be *i.i.d* across years and hospitals. I pool all the observations together and estimate the LPM with OLS. The IV estimation is a standard TSLS approach.

2.7 Empirical Results

Taking non-adoption as the baseline response, I regress hospitals' choice on a number of variables, including both vendor and hospital characteristics. In different specifications, I include different variables in Table 2.7. Other control variables contain the bed size, profit status, number of competitors in the local market, ratio of adoptions by neighboring hospitals, whether a hospital is an academic medical center, and percentage of medicare and medicaid discharges. Both hospital and year fixed effects are included in all specifications, and the standard errors are clustered at the hospital level. Table 2.8 shows the results of the estimation based on all affiliated hospitals. The first column lists the variables of interest.[3] The second through the last column report the estimated coefficients in different specifications. The first row reports the coefficients for the indicator of whether a particular vendor is chosen by the hospital in the previous year. They are all positively significant and close to unity, which implies the previous choice plays an important role in forming the current decision. Hospitals will not easily switch vendors once they install the software. The estimated coefficient for the market leader dummy is not significant in all specifications, implying that being a local leading provider of EMR merely has any advantage. The impact from choosing a system-dominant vendor is positive and highly significant. The estimates for this variable have similar values when I control for the size of the hospital and parent system. The coefficients for the interaction between the market leader dummy and the system size is positive at the 1% level of significance. It is suggested the more "spread-out" or the larger a hospital chain is, the more likely for its member hospitals to coordinate with the local market. The size of the parent system has a greater impact in affecting the hospital's choice of vendors than that of the hospital itself since all the coefficients associated with the large hospital dummy are not significant.

I further run the same regressions on the sample of different types of hospital systems. In particular, I examine whether the behavior of multi-region hospital systems differs from that of hospital chains locating in one market. The left panel in table 2.9 presents the results for

[2]The estimates from ivprobit and controlled function are available upon request.

[3]The estimates of other control variables are omitted here due to limited space, but they are available upon request.

the sample of multi-region hospital chains. The estimated coefficients in the first row suggest that the adoption choice in the previous period plays an important role in the current decision. The estimates of the indicator for the system-dominant vendor are significant and positive in all specifications. The coefficients for the market leader dummy alone are insignificant, but its interactions with system size measures are positively significant, showing a similar effect on the entire sample. The right panel reports the coefficients of the regressions based on the sample of single-region hospital systems. In the last column, the coefficients for the market leader indicator is positive at the 5% level of significance but its interaction with the system size is significantly negative, suggesting larger systems are more likely to deviate from the local market. In all, following the choice of the parent system brings in higher profits from adoption, compared with that from choosing the most popular products in the local market. The impacts from product characteristics do not vary between large and small hospitals.

As indicated above, there may exist confounding elements at the market level that I aim to separate in the estimation. The first set of instruments describes the market condition of the associated outside markets, which relies on the identifying assumption of no spill-over across markets. If inter-market dependency cannot be ruled out, the second set of instruments—the adoption status of single-region hospital systems in the associated outside markets—could be applied to further correct this problem. However, it does not come with no price given the possibility of being weak instruments. The estimation becomes more complicated with the inclusion of the interaction terms. TSLS is one of the most "harmless" options, which is convenient to implement and whose results are easy to interpret. It is worthwhile to point out the potential bias due to endogeneity. I expect the unobserved components are positively correlated with both the choice probability and the endogenous variables. In the appendix, I provide a simple proof of potential bias due to omitted variables. It is suggested that the estimated coefficient for the market leader dummy is upper biased while the effect is uncertain for interaction terms without the correction of instruments.

Table 2.10 presents the results of the IV estimation using the first set of instruments. In this case, I use the markets with affiliated hospitals at least one of whose competitors belong to a multi-region hospital system. This reduces the sample by about one third. The left panel provides the estimates of the non-IV estimation based on this sample while the right panel shows the results from the IV regression. The estimated coefficients of both indicators—whether a particular vendor is system-dominant and whether it is previously-chosen—are consistently significant in both non-IV and IV estimations. Moreover, both the sign and magnitude are very robust across all specifications. However, the estimated coefficient for the market leader dummy is insignificant, regardless of the usage of instruments. But the estimates of the interaction terms between the local leader dummy and the parent system size become negatively significant after the IV is applied, suggesting that large hospital chains are more reluctant to coordinate with the local market. The estimates for the interaction terms involving the system-dominance indicator are always positively significant. The magnitude of these coefficients become even larger in the IV estimation.

The next table (Table 2.11) shows the results from the same specification using the second set of instruments. I use the adoption status of the single-region hospital systems locating in the outside associated markets to rule out potential interdependence across markets. The sample is further restricted to markets including affiliated hospitals at least one competitor of which belongs to a multi-region hospital system whose majority members locate in a market with at least one single-region hospital systems. All the IV estimations pass the tests for weak instruments, suggesting evidence of sufficient relevance between the instruments and endogenous variables. The previously-chosen vendor is still a very strong predictor for the current choice, as shown in the first row. Moreover, the impact has been increasing in the IV estimation. The estimates for the system-dominant vendor dummy show a similar pattern. The estimated coefficients become greater by more than 30% when instruments are adopted. What is most interesting is that the IV estimates for the market leader dummy become negative at the 5% level of significance and it is quite robust in all the IV specifications. It is suggested that being a local market leader will decrease the probability of being adopted by about 0.16-0.25 units. This may uncover hospital chains' tendency to differentiate from the local market. The results imply internal consistency is much more attractive to affiliated hospitals, and these hospitals seem unable to enjoy the complementarities from the local market. If a member hospital has to choose between the system-dominant vendor and the most popular local supplier, being consolidated to the parent system has at least two advantage: One is to enjoy economies of scale and another is less competition pressure in the local market. It may reflect the strategic component in the choice of EMR vendors for hospital systems, which is somewhat consistent with the findings in Miller and Tucker (2014).

2.8 Conclusion and Future Work

The goal of this paper is to understand hospital systems' incentive in choosing EMR vendors. When an affiliated hospital faces a variety of vendors, it may have to consider choosing the most popular vendor in the local market or the major Health IT provider for the parent system. In this paper, I use a simple discrete-choice model to examine how hospital systems assess the value of these two factors. The result of the analysis has some potentially important policy implication. If hospitals are more likely to use the product as mostly-adopted by local hospitals, it reflects the hospital's willingness to coordinate with the outside market. If instead the vendor dominating the parent system is much more preferable to the local market leader, it may create barriers to market integration for the application of EMR. With further understanding of the driving forces underlying their behavior, policy makers will be better positioned in promoting the adoption of Health IT.

The preliminary analysis suggests that when the local market leader is distinct from the system-dominant vendor, the number of hospitals following the choice of the parent system is much greater than that of those choosing the largest EMR vendor in the local area. The difference is even more significant for the sample of single-region hospital systems. The result is consistent with that based on more rigorous analysis. The estimates after controlling a rich set of variables implies little benefits from adopting the market-leading technology, compared with the consistent evidence of significant gains from choosing the system-dominant vendor. I use two types of instruments to correct the potential endogeneity issue. The IV estimation strengthens the positive impacts from being integrated with the parent system. However, it even flips the sign to negative for the estimated effect of choosing the local largest vendor in some specifications. One possible explanation is that by following the choice of the parent system, the member could achieve high levels of economies of scale. Moreover, it may also reflect hospitals' potential reluctance to coordinate with outside market, which could possibly go against the government's intention to build the Nationwide Health Information Network.

The estimation of this model involves some complication given the potential endogeneity issue. I apply the LPM as it is one of the most "harmless" methods according to the literature and as it is easy to implement and interpret. I consider exploring other estimation approaches for robustness check in addition to ivprobit and control functions. Furthermore, the model capturing the adoption behavior is built on a simple static framework. Taking account that the adoption decision could involve forward-looking behavior, I also consider introducing dynamic modelling with richer structure. But the model setting requires more careful thoughts given that the adoption decision for affiliated hospitals is such a complex process.

Table 2.1: Average percentage distribution for all hospital systems ($N = 4185$)

Choice	Percentage
Market-leading	9.27% (0.29)
System-dominant	69.89% (0.4587)
Neither	20.84% (0.4061)

Note: The sample used here is conditioning on market-leading and system-dominant being different. Standard errors are in parentheses.

Table 2.2: Average percentage distribution for multi-region hospital systems ($N = 3556$)

Choice	Percentage
Market-leading	10.15%
	(0.302)
System-dominant	67.38%
	(0.4688)
Neither	22.47%
	(0.4174)

Note: The sample used here is conditioning on market-leading and system-dominant being different. Standard errors are in parentheses.

Table 2.3: Average percentage distribution for single-region hospital systems ($N = 629$)

Choice	Percentage
Market-leading	4.29% (0.2027)
System-dominant	84.1% (0.3657)
Neither	11.61% (0.3203)

Note: The sample used here is conditioning on market-leading and system-dominant being different. Standard errors are in parentheses.

Table 2.4: Top 11 EMR vendors and their market share among in 2006

Vendor	Market share
Healthland	1.91%
EPIC	2%
Healthcare Mngt. Systems (HMS)	2.31%
GE Healthcare	3.33%
Eclipsys	3.87%
CPSI	5.69%
Self-developed (SD)	7.69%
Cerner	11.73%
Siemens	11.87%
McKessons	12.67%
Meditech	28.98%
sum	92.04%

Note: The calculations are based on the sample.

Table 2.5: Summary statistics in the sample of 2006

Variable	Obs	Mean	Std. Dev.	Min	Max
EMR adoption for all hospitals	4560	0.49	0.50	0	1
Number of hospitals in a HSA	4560	4.95	6.49	1	90
Ratio of competitors adopting EMR	4560	0.37	0.37	0	1
EMR adoption for affiliated hospitals	2500	0.56	0.50	0	1
Number of affiliated hospitals in a HSA	2500	2.71	4.62	0	57
% for affiliated hospitals to use system-dominant products	1407	0.85	0.36	0	1
% for affiliated hospitals to use market-leading products	1407	0.69	0.46	0	1
Number of beds for affiliated hospitals	2500	187.31	194.81	6	2205
For-profit affiliated hospital	2500	0.28	0.45	0	1
Not-for-profit affiliated hospital	2500	0.61	0.49	0	1
Academic medical center	2500	0.071	0.26	0	1
% of Medicare discharge for affiliated	2500	0.47	0.16	0	1
% of Medicaid discharge for affiliated	2500	0.16	0.11	0	1

Table 2.6: Summary statistics in the sample of 2009

Variable	Obs	Mean	Std. Dev.	Min	Max
EMR adoption for all hospitals	4560	0.84	0.37	0	1
Number of hospitals in a HSA	4560	4.95	6.49	1	90
Ratio of competitors adopting EMR	4560	0.67	0.41	0	1
EMR adoption for affiliated hospitals	2591	0.89	0.31	0	1
Number of affiliated hospitals in a HSA	2591	2.81	4.66	0	56
% for affiliated hospitals to use system-dominant products	2318	0.78	0.41	0	1
% for affiliated hospitals to use market-leading products	2318	0.60	0.49	0	1
Number of beds for affiliated hospitals	2591	190	199.63	1	2249
For-profit affiliated hospital	2591	0.27	0.44	0	1
Not-for-profit affiliated hospital	2591	0.63	0.48	0	1
Academic medical center	2591	0.07	0.26	0	1
% of Medicare discharge for affiliated	2591	0.47	0.15	0	0.98
% of Medicaid discharge for affiliated	2591	0.17	0.11	0	0.81

Table 2.7: Variables in X_{it}^{j}

Variables
Indicator of whether vendor j is system-dominant
Indicator of whether vendor j is market-leading
Indicator of whether vendor j is chosen by hospital i at $t-1$
Indicator of whether vendor j is system-dominant $\times$ indicator of large hospital
Indicator of whether vendor j is market-leading $\times$ indicator of large hospital
Indicator of whether vendor j is system-dominant $\times$ # of HSAs the hospital system locate
Indicator of whether vendor j is market-leading $\times$ # of HSAs the hospital system locate
Indicator of whether vendor j is system-dominant $\times$ # of members in the hospital system
Indicator of whether vendor j is market-leading $\times$ # of members in the hospital system

Note: The hospital system in the last four rows refer to the system hospital i affiliates to.

Table 2.8: LPM Estimation based on all affiliated hospitals

	All hospital systems		
same_as_previous	0.889***	0.889***	0.888***
	(0.00485)	(0.00482)	(0.00486)
mktlead	0.000804	-0.00188	-0.00246
	(0.00165)	(0.00239)	(0.00246)
sysdom	0.0807***	0.0799***	0.0850***
	(0.00442)	(0.00552)	(0.00563)
mktlead×big		-0.0000341	0.000499
		(0.00347)	(0.00349)
sysdom×big		-0.00742	-0.00873
		(0.00562)	(0.00566)
mktlead×nmb		0.0000880***	
		(0.0000337)	
sysdom×nmb		0.000235**	
		(0.000105)	
mktlead×nhsa			0.000134***
			(0.0000512)
sysdom×nhsa			0.000154
			(0.000141)
adj. R-sq	0.781	0.781	0.781
N	124668	124668	124668

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 2.9: LPM Estimation based on different types of affiliated hospitals

	Multi-region hospital systems			Single-region hospital systems	
same_as_previous	0.886***	0.887***	0.886***	0.904***	0.904***
	(0.00519)	(0.00516)	(0.00519)	(0.0144)	(0.0142)
mktlead	-0.0000172	-0.00362	-0.00414	0.00540	0.0216**
	(0.00181)	(0.00266)	(0.00276)	(0.00384)	(0.00855)
sysdom	0.0829***	0.0806***	0.0864***	0.0606***	0.0750***
	(0.00467)	(0.00589)	(0.00602)	(0.0141)	(0.0197)
mktlead×big		-0.00157	-0.000907		-0.000934
		(0.00382)	(0.00382)		(0.00716)
sysdom×big		-0.00253	-0.00380		-0.0178
		(0.00626)	(0.00627)		(0.0122)
mktlead×nmb		0.000109***			-0.00560**
		(0.0000363)			(0.00267)
sysdom×nmb		0.000211*			-0.00187
		(0.000113)			(0.00363)
mktlead×nhsa			0.000159***		
			(0.0000550)		
sysdom×nhsa			0.000107		
			(0.000151)		
adj. R-sq	0.773	0.774	0.774	0.822	0.823
N	103837	103837	103837	20831	20831

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 2.10: IV Estimation using outside markets as instruments

	Non-IV estimation			IV estimation		
same_as_previous	0.887***	0.888***	0.887***	0.888***	0.890***	0.888***
	(0.00601)	(0.00596)	(0.00599)	(0.00297)	(0.00302)	(0.00295)
mktlead	-0.00131	-0.00174	-0.00238	-0.00977	0.00828	0.00249
	(0.00197)	(0.00285)	(0.00295)	(0.0115)	(0.0169)	(0.0160)
sysdom	0.0791***	0.0704***	0.0755***	0.0800***	0.0667***	0.0744***
	(0.00539)	(0.00658)	(0.00673)	(0.00240)	(0.00489)	(0.00441)
mktlead×big		-0.00203	-0.00171		0.0239	0.0226
		(0.00424)	(0.00430)		(0.0164)	(0.0163)
sysdom×big		-0.000266	-0.000748		-0.00545	-0.00610
		(0.00681)	(0.00688)		(0.00567)	(0.00566)
mktlead×nmb		0.0000522			-0.000849**	
		(0.0000384)			(0.000390)	
sysdom×nmb		0.000447***			0.000550***	
		(0.000137)			(0.0000876)	
mktlead×nhsa			0.0000943			-0.000875**
			(0.0000582)			(0.000437)
sysdom×nhsa			0.000423**			0.000489***
			(0.000187)			(0.000107)
N	93559	93559	93559	93559	93559	93559

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 2.11: IV Estimation using single-region hospital systems as instruments

	Non-IV estimation			IV estimation		
same_as_previous	0.887***	0.889***	0.888***	0.912***	0.914***	0.912***
	(0.00667)	(0.00661)	(0.00664)	(0.0122)	(0.0129)	(0.0130)
mktlead	-0.000960	0.000497	-0.000203	-0.165**	-0.245**	-0.252**
	(0.00228)	(0.00334)	(0.00344)	(0.0792)	(0.105)	(0.116)
sysdom	0.0797***	0.0709***	0.0769***	0.0950***	0.111***	0.119***
	(0.00606)	(0.00736)	(0.00750)	(0.00778)	(0.0182)	(0.0204)
mktlead×big		-0.00282	-0.00260		0.0213	0.0333
		(0.00474)	(0.00480)		(0.0853)	(0.0855)
sysdom×big		-0.00349	-0.00407		-0.00556	-0.0117
		(0.00758)	(0.00768)		(0.0225)	(0.0227)
mktlead×nmb		0.00000671			0.00180	
		(0.0000425)			(0.00119)	
sysdom×nmb		0.000536***			-0.000103	
		(0.000166)			(0.000304)	
mktlead×nhsa			0.0000352			0.00305
			(0.0000656)			(0.00237)
sysdom×nhsa			0.000515**			-0.000415
			(0.000229)			(0.000458)
N	74476	74476	74476	74476	74476	74476

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

CHAPTER 3

3.1 Introduction

This paper examines whether the response to the Medicare payment reform in 2007 differed between hospitals with and without Electronic Medical Records (EMR). In particular, we focus on the payment differences arising under the acute care hospital inpatient prospective payment system (IPPS). The recent payment reform aims to account more fully for the severity level of patient conditions. There is mixed evidence on the impact of the payment reform on medical billings and collections. One line of evidence suggests that EMR leads to increases in these parameters. By comparing the behavior of hospitals with and without Health IT, we explore whether the application of the new technology brings in more vulnerabilities under the new payment structure. The result of this study can impact our understanding of the mechanisms behind how health IT might affect hospital costs under different reimbursement environments.

The theory that EMR facilitates upcoding dated back to an article published in September 2012 from New York Times, which documented multiple cases of sharp rises in the higher-pay claims right after the providers adopted EMR. By analyzing the data between 2006 and 2010, the article indicated that hospitals receiving the meaningful-use reimbursement for Health IT experienced a 47% increase in Medicare payment, compared with the 32% rise in the hospitals that hadn't received any incentive payment. There are three major explanations for why overall coding could increase due to Health IT: better documentation; sicker patients; and physician gaming the system. We try to examine whether the technology inflated coding. We obtain identification from a recent payment reform. In the shift to the new payment system, the Center for Medicare and Medicaid Services (CMS) created finer categorization to measure the differences in patent severity, which could further affect the role played by EMR in coding practices.

We document hospitals' differing response to this reform in the presence of Health IT. It is deemed imperative for policy making in many aspects particularly since the adoption of EMR continues to expand. If the use of more remunerative billing codes is accelerating partly due to the transition to EMR, the government should take appropriate steps to ensure the proper use of the new technology and avoid billions of overpayment. With the information on how Health IT affects coding process, the policy makers are able to provide more guidance to address the vulnerabilities in EMR. After all, the risk of misuse should not impede the diffusion of this technology given its potential benefits. The goal of this project is to fill gaps in this evidence base and provide sound empirical support for future policy.

With the 100% Medicare Claims data on the inpatient sector, we construct two outcome measures—the Case-mix Index (CMI) and the fraction of patients assigned to higher-pay codes—and compare the differences between hospitals that have and have not adopted Health IT. Using the difference-in-difference estimation technique, we found the responses vary between surgical and medical codes before and after the reform, which could potentially reveal some information on how the technology influences the coding practices.

The rest of the paper proceeds as follows: Section 2 presents literature related to this topic. Section 3 provides basic information and institution background about the payment reform and EMR. Section 4 describes the datasets applied in this study. Section 5 describes the estimation strategy. Section 6 shows the estimation results. The last section concludes and points out the future directions.

3.2 Related Literature

There has not been much literature about the impacts of Health IT on the medical billing in spite of the intensive discussion. Li, in her job market paper (2013), measured the effect of EMR adoption on medical coding and billing practice. She used the multi-state inpatient discharge data to examine the relationship and found that the share of patients coded to a higher-pay DRG increased significantly after EMR adoption. A more recent study by Adler-Milstein and Jha (2014) used longitudinal data on EMR adoption to examine whether the institutions with Health IT experienced greater increase in patient acuity and payment from Medicare. They identified a group of adopted hospitals and selected up to three control institutions for each treated hospital. Focusing on the comparison on two outcome measures, the Case-mix index and payment per discharge, they found no evidence of higher Medicare reimbursement for hospitals with EMR. Li's job market paper only covered the period before the policy reform while the study by Adler-Milstein and Jha analyzed the post-reform impacts. Our paper is distinct from these studies in the sense that we assess how Health IT influences hospitals practices in response to the reform by examining both the pre- and post-reform periods. Furthermore, we try to uncover the potential mechanism driving this outcome.

A large strand of literature documented hospitals behavior in the presence of the PPS, which could potentially lead to a phenomenon called DRG creep or upcoding, i.e., healthcare providers intentionally group patients into more resource-intensive DRG classifications in order to increase hospital income. Coulam and Gaumer (1991) reviewed a large body of work that had contributed to the understanding of the effectiveness and the potential outcomes of the program. They summarized some evidence of upcoding, which mostly came from the analysis based on the Case-mix index. A well-cited study by Dafney (2005) used a different measure—the fraction of patients assigned to the higher-pay codes—to examine hospitals' response to diagnosis-related price changes. By exploiting an exogenous policy reform in 1988, she found hospitals had incentive to upcode patients and the effect was strongest among for-profit hospitals, which is consistent with the findings in an earlier study by Silverman and Skinner (2004). Such upcoding practices were also documented in skilled nursing facilities and the outpatient sector (Brunt (2011); Bowblis and Brunt (2013)). Our paper also contributes to this strand of literature as we take account of the role played by Health IT in affecting the hospitals' practices in response to payment reforms.

3.3 Background

3.3.1 Policy Change

In 1983, the Health Care Financing Administration (HCFA, now the Centers for Medicare and Medicaid Services (CMS)) implemented diagnosis related groups (DRGs) for the IPPS. One DRG is assigned for each patient stay according to the principal diagnosis, additional diagnoses, the principal procedure and additional procedures, age, sex and discharge status. Each DRG has a payment weight assigned to it, based on the average resources used to treat Medicare patients in that DRG. In the DRG framework, each hospital is paid a flat rate per episode of inpatient care. A hospital makes profits if it treats a patient with less spending than the DRGs payment. The DRGs prospective payment system (PPS) aims to reward efficiency and contain the growing expenditure in the health care sector.

The process of forming DRGs starts by dividing all possible diagnoses into one of the 25 Major Diagnostic Categories (MDC), each of which generally corresponds to a single organ system or etiology, except for some patient groups that are extremely resource-intensive. Those exceptional cases such as organ transplants, bone marrow transplants and tracheostomy cases, called pre-MDC DRGs, are put into a separate DRG group before MDC assignment based on the operating room (O.R.) procedure rather than principal diagnosis. Once the MDCs are defined, they will be determined as surgical DRGs or medical DRGs according to whether they involve O.R. procedure or additional patient characteristics that will impact the consumption of hospital resources. Some surgical and medical DRGs are further identified with the presence of complications or comorbidities, if their secondary diagnosis causes substantially increased hospital resource use. After the patient's information on diagnoses, procedure, discharge status and demographic information is entered into the Medicare claims processing systems and goes through a series of automated screens, an appropriate DRG will be classified by the Medicare GROUPER software system.

In August 2007, the CMS announced its final rule on the payment reform—shifting to the new Medicare Severity DRGs (MS-DRGs)—for inpatient hospital services, with the goal to improve the payment accuracy under the IPPS. Particularly, the reform restructured the inpatient DRGs in order to better reflect the severity level of patient conditions. CMS expanded the new version of DRGs by replacing the old 538 DRGs with 745 new severity-adjusted DRGs. In the old version, there were some instances of a pair of DRGs sharing a common primary diagnosis, one capturing some level of Complication and Comorbidities (CC) and another for no CC severity level. In the new version, the CMS reviewed some concepts of what diagnoses to be considered as complications or comorbidities and used two levels to replace the original single level: (1) Complication and Comorbidities (CC), and (2) Major Complications and Comorbidities (MCC). This significant refinement was to create a three-tiered system with greater granularity in measuring differences in patient severity.

The MS-DRGs resequenced the group number and readjusted the weights for all diagnoses. Table 3.1 displays some examples of the transformation in the reform. Of considerable relevance and importance, and particularly germane to our paper, MS-DRGs were intended to provide more appropriate reimbursements for care conducted at tertiary care hospitals, relative to specialty hospitals. Many of the latter focus on elective or semi-elective surgical interventions (cardiac, orthopedic) on patients with less severe conditions or underlying morbidities. Higher risk individuals are preferentially targeted to tertiary care centers. Since the reform was structured to be revenue neutral, changes in overall surgical coding would be expected, albeit of uncertain direction and magnitude. Less clear would be effects on coding for medical conditions.

3.3.2 EMR and Coding

EMR were invented in 1970s, but the acceptance to this technology had been very slow until recent years. In 2009, the American Recovery and Reinvestment Act (ARRA) has provided $27 billion to promote Health Information Technology (HIT), in particular to encourage the adoption of EMR. It is the first substantial commitment of federal resources to support the adoption of EMR and creates a strong push in the diffusion of HIT. As the cornerstone of the Affordable Care Act in improving quality and lowering cost, EMR serves functions that paper records cannot deliver. According to the Healthcare Information and Management Systems Society (HIMSS), a solid EMR foundation should include the following key components: Clinical Data Repository (CDR), Clinical Decision Support Capabilities (CDS), and Computerized Physician/Provider Order Entry (CPOE). CDR is essentially a centralized database that collects, stores, accesses and reports health information. It is the backbone of the entire system. CDS assists clinicians in decision-making tasks namely determining the diagnosis or setting treatment plans. CPOE is a more advanced type of electronic prescribing. It can link to the adverse drug event (ADE) system to avoid potential medication errors.

The adoption of EMR has substantially altered the practices that could potentially boost up coding. The seemingly inflated billing could be possibly due to the fact that the providers had previously downcoded during the era of paper records. In the past, clinicians have to provide sufficient details in order to justify the reimbursement while with digitized systems a few clicks can readily produce comprehensive medical records. EMR facilitates the documentation and billing process such that it is deemed to improve accuracy in the billing cycle.

Certain functionalities of EMR, such as generic pick lists, preloaded macro, autofill, and copy-and-paste, especially make documentation of patient severity, including underlying conditions, more automatic and less labor intense than is the case with paper records. This feature is desirable if the consequence is more complete and accurate documentation. One particular example is charting by exception. EMR makes it very convenient to create a complete list of examinations. Physicians may rush through their EMR charting without deleting the procedures that have not been performed. Moreover, there also could be financial pressure for physicians to upcode, for which the adoption of EMR easies the path for inflated billing.

As described in anonymous articles, physicians have incentive to game the system with Health IT. There are two types of misuse of EMR that can inflate the Medicare billing. The first one, called cloning, occurs when the physician uses the cut-and-paste technique in a medical record, inadvertently neglecting to delete some items, therefore making it appear as though the same medical procedures were performed at a later date. The cloning document can also be created by duplicating the same service that was never rendered for multiple patients. The second one, catching more attention, is a practice known as upcoding. For instance during every office visit/encounter, the physician needs to select the appropriate Evaluation and Management (E&M) codes for Medicare reimbursement, which reflect the range of care delivered and the time it takes. Doctors could possibly assign patients to the higher-level codes—the practice of charging for more extensive and costly services than delivered—facilitated by the application of Health IT.

While these practices have gained attention, there is a powerful counterbalance that is often ignored or dismissed. Physicians and hospitals will sustain huge financial and reputational losses if they inappropriately upcode, and are audited by Medicare. Penalties include repayment and triple damages, in many if not most cases. High profile examples of aggregate penalties in the hundreds of millions of dollars for a single hospital or system are well known. As a consequence, those entities, and even more so the individuals in the "back office" who generate the claims, strive for precision in the billing and coding process. In fact, a common phenomenon is under-coding, as a hedge against potential Medicare fraud. Since EMR has the realized potential to increase accuracy of records, it should not be surprising if Health IT augments claims in some situations.

3.4 Data

The data is constructed by pooling information from various sources. The first primary dataset comes from the Medicare Provider Analysis and Review (MedPAR) File which contains the information about final action stay records of Medicare beneficiaries in Medicare certified inpatient hospitals and Skilled Nursing Facilities (SNF). Each MedPAR record represents a stay, which is an accumulation of claims from the data of admission to an inpatient hospital, or to a skilled nursing facility. A stay record may represent one claim or multiple claims. We use the data of 100% Medicare beneficiaries, which cover more than 16 million observations per year. The claims data provides us sufficient information to construct the outcome measures.

The technology data comes from the Healthcare Information and Management Systems Society (HIMSS) Analytics Database, which is the most comprehensive national source of hospital information technology (HIT) adoption data. We use the Medicare provider number to connect to the claims data. The database covers the majority of U.S. hospitals and includes market share and purchasing plan data for over 90 software applications and technologies. It is an annual survey recording the time and the choice of a hospital's adoption decision. More specifically, the dataset contains the information about the identity of the product supplier, enabling a more realistic network definition.

There is no consensus on how to define adoption of EMR for a hospital. Jha et al. (2009)

used a very comprehensive definition. From a list of 32 potential functionalities of an inpatient electronic health record, they asked an expert panel to define the functionalities that constitute a basic and comprehensive electronic system respectively. Miller and Tucker (2009) measured EMR adoption by whether a hospital is installing or has installed the enterprise EMR system. In our paper, a hospital is defined to adopt EMR if CDS is live and operational in the hospital. CDS, as one of the key components of the entire IT system, is the part involving physicians most. Since we pay special attention to how Health IT alters physicians' behavior in response to the payment reform, the adoption status of this component is most relevant to identify the effect.

We complement the HIT data with the American Hospital Association (AHA) Annual Survey, using the Medicare provider number and geographic information to perform the linkage. The AHA data includes a rich set of hospital-specific features such as number of beds, system affiliation, profit status and etc. Table 3.2 and 3.3 provide summary statistics for the main variables before and after the reform. The first row shows the statistics for the fraction of patients classified to the higher-pay category. Note that there are instances of a pair or trio of DRGs that share a common diagnosis but represent different levels of severity. For each of this pair or trio, we compute how many of the patients having the same primary diagnosis were assigned to the group with at least some complications or comorbidities. There were 76.7% of patients with some presence of CC before the reform but the ratio dropped to 56.9% after the reform. The non-trivial decrease arises from the fact that the MS-DRGs were estimated to produce a much higher percentage of discharge assigned to the lowest severity level, according to a study by Rand (2007) about an evaluation of alternative DRG systems. Without such an adjustment in the MS-DRGs GROUPER software, the aggregate payment for inpatient hospital services is predicted to increase significantly. In 2006, 35.5% of the hospitals adopted EMR and the number went up to 57.2% after the reform. The Case-mix index as well as other hospital characters had not changed much during the sample period.

3.5 Estimation Method

Our paper attempts to document hospitals' differing response to a recent Medicare payment reform in the presence of Health IT. We estimate the relation based on the Case-mix index and the fraction of upcoded patients, measured by the fraction of patients assigned to the class with at least some presence of CC for each DRG pair or trio. The Case-mix index, calculated as the volume-weighted average DRG weight, is stated to give appropriate payment implications according to the literature. It is an aggregate measure of a group of DRGs, accounting for both the DRG weights and the corresponding patient volume. However, in our context of the payment reform, it remains uncertain how much of the change in Case-mix index originates from the requirements of more intensive-resource usage, or from the readjustment of the DRG weights, or from the interaction between both, or from a change in the distribution of DRGs for admitted patients. Thus, we also use another measure—the fraction of upcoded patients—to estimate the difference. As a result, more emphasis is placed on the primary diagnoses enabling the calculation of the fraction variables, i.e., the DRG pairs that weren't changed and the DRGs that were expanded from one to two, one to three and two to three tiers in the reform.

We compare the difference in the Case-mix index and the fraction between hospitals with and without EMR. The Case-mix index can be calculated by summing the DRGs weights, weighted by the corresponding discharge. Each hospital per year involves one index. For each DRG pair or trio, we compute the fraction of upcoded patients and pick out the DRGs for which the fraction is higher in hospitals with EMR at least at the 10% level of significance. We identify the difference under a difference-in-difference framework. The dependent variable is one of the outcome measures and the independent variables contain the key variables of interest—whether the hospital has adopted EMR and whether the outcome measure is associated with medical codes—as well as other control variables. We estimate the regression for both the pre- and post- reform periods to see how different the response of hospitals could be. All specifications include hospital fixed effects, year dummies and the indicators for each primary diagnosis.

3.6 Results

For each primary diagnosis, we compare the fraction of upcoded patients between hospitals with and without EMR during 2008 to 2010. There are 186 of such DRGs we consider, and more than 70% of them are medical codes. Table 3.4 lists the information of several consistently higher-coded diagnoses in the presence of health IT. The first column shows the code for each DRG, the second shows the title, the next shows the difference in the fraction between two types of hospitals and the last reports the corresponding DRG weight. The first eight rows represent medical diagnoses and the remaining are surgical. We find the patients in medical DRGs were more often assigned to the top codes than those in surgical diagnoses. Table 3.5 shows the change of the Case-mix index over time. The upper panel presents the Case-mix index based on all DRGs, the middle one shows the indices based on surgical DRGs and the lower one is for medical diagnoses. The first column shows the Case-mix index over all hospitals and the last two report the average indices across hospitals with and without Health IT respectively. In the upper panel, the Case-mix index had not changed much over the years, with an accumulating increase by 7%. When we compare the indices across hospitals, the Case-mix index for hospitals with EMR is approximately 10% higher than that for hospitals without EMR. However, we do not see such a difference for the Case-mix index based on surgical and medical codes. Moreover, the indices for surgical codes are more than twice higher than those for medical codes, reflecting more clinical complexity and greater needs for resources in treating surgical population.

Table 3.6 reports the estimated coefficients from the regression of the Case-mix index on a rich set of variables capturing hospitals' characteristics. The key variable of interest is the indicator of whether the hospital has adopted EMR. A hospital is defined to have adopted EMR if the component of CDS is live and operational within the organization. We only report the estimated coefficient of the dummy for adoption. The upper panel presents the results based on the periods prior to the reform and the lower panel shows the results for the post-reform estimation. We repeat the estimation for the Case-mix index on surgical and medical codes respectively, as shown in the last two columns. There is no significant difference in the general Case-mix index between hospitals with and without EMR, regardless of the time frame. It is consistent with the findings by Adler-Milstein and Jha (2014) who detected no evidence of upcoding resulting from the adoption of Health IT. The results become more interesting when we separately examine the Case-mix indices for surgical and medical codes. Prior to the reform, the surgical index is significantly higher among hospitals with Health IT but the sign flipped after the reform. Table 3.7 presents the coefficients of regressing the fraction of upcoded patients on the same set of control variables. We additionally include the indicator of medical diagnosis as well as its interaction with the adoption dummy, aiming to capture the varying effect of adoption on different types of diagnoses. As indicated in the last row, the estimates of the same interaction term has opposite signs based on the data after the reform. It suggests that before the reform, there was a larger fraction of surgical patients, than medical patients, assigned to the more resource-intensive category in hospitals with Health IT, but the outcome reversed in the post-reform period.

3.7 Conclusion

Converting to digitized medical records has, to a large extent, altered the operational practices of health care providers. Billions of dollars have been invested on this technology without clear evidence of effective implementation. This paper examines one of the unintended outcomes of Health IT—whether the application of this technology brings in extra vulnerabilities to the Medicare payment system. In particular, we try to explore the mechanism by which Health IT influences hospital behavior in response to the transition to the MS-DRGs system. By making use of this payment reform, we attempt to further understand how EMR affects the documentation and billing process. We find, in general, the adoption makes no significant difference in billing, but the impact of Health IT on medical/surgical diagnoses reversed after the reform.

Table 3.1: Examples of changes in DRGs

Expand from 1 to 2

	Code	Title	Weight
Old	9	Spinal disorders & injuries	1.4045
New	52	Spinal disorders & injuries w CC/MCC	1.4329
	53	Spinal disorders & injuries w/o CC/MCC	1.1172

Expand from 1 to 3

	Code	Title	Weight
Old	127	Heart failure & shock	1.0345
New	291	Heart failure & shock w MCC	1.2585
	292	Heart failure & shock w CC	1.0134
	293	Heart failure & shock w/o CC/MCC	0.8765

Expand from 2 to 3

	Code	Title	Weight
Old	531	Spinal procedures w/ CC	3.1279
	532	Spinal procedures w/o CC	1.4195
New	28	Spinal procedures w MCC	4.2339
	29	Spinal procedures w CC	2.8356
	30	Spinal procedures w/o CC/MCC	1.7617

Table 3.2: Summary statistics for key variables before the reform

Variable	Obs	Mean	Std. Dev.	Min	Max
Fraction	212564	0.767	0.29	0	1
Case-mix index	4437	1.236	0.315	0.6042641	4.47
Adopt	4437	0.355	0.479	0	1
Bed size	4437	168.58	183.031	6	2205
Whether it is a teaching school	4437	0.066	0.248	0	1
Total number of outpatient visits	4437	126564	188857.4	0	3109854
Total number of admissions	4437	7491.165	9253.582	12	107845
Total number of full time physicians and dentists	4437	15.036	70.159	0	1674
Percentage of Medicare discharge	4437	0.488	0.15	0	1
Percentage of Medicaid discharge	4437	0.163	0.103	0	0.89
Whether it is a for-profit hospital	4437	0.167	0.373	0	1
Whether it is a not-for-profit hospital	4437	0.596	0.491	0	1
Total number of births	4437	868.975	1383.257	0	18598

Table 3.3: Summary statistics for key variables after the reform

Variable	Obs	Mean	Std. Dev.	Min	Max
Fraction	207285	0.569	0.358	0	1
Case-mix index	4354	1.291	0.357	0.572	5.37
Adopt	4354	0.572	0.495	0	1
Bed size	4354	171.085	189.283	6	2204
Whether it is a teaching school	4354	0.067	0.25	0	1
Total number of outpatient visits	4354	133721.8	198696.1	0	3341664
Total number of admissions	4354	7650.619	9615.374	9	111203
Total number of full time physicians and dentists	4354	18.149	85.358	0	1886
Percentage of Medicare discharge	4354	0.489	0.143	0	1
Percentage of Medicaid discharge	4354	0.166	0.101	0	0.749
Whether it is a for-profit hospital	4354	0.159	0.366	0	1
Whether it is a not-for-profit hospital	4354	0.602	0.49	0	1
Total number of births	4354	875.546	1378.412	0	17203

Table 3.4: Consistently higher-coded diagnoses

DRG code	Title	Difference	DRG weight
314	Other circulatory system diagnoses w MCC	0.0836	1.5606
315	Other circulatory system diagnoses w CC	0.0836	1.172
316	Other circulatory system diagnoses w/o CC/MCC	0.0836	0.9075
542	Pathological fractures & musculoskelet & conn tiss malig w MCC	0.0621	1.4877
543	Pathological fractures & musculoskelet & conn tiss malig w CC	0.0621	1.1151
544	Pathological fractures & musculoskelet & conn tiss malig w/o CC/MCC	0.0621	0.9395
871	Septicemia w/o MV 96+ hours w MCC	0.0801	1.7484
872	Septicemia w/o MV 96+ hours w/o MCC	0.0801	1.3783
488	Knee procedures w/o pdx of infection w CC/MCC	0.0579143	1.6584
489	Knee procedures w/o pdx of infection w/o CC/MCC	0.0579143	1.4512

Table 3.5: Case-mix index for different groups of DRGs

All DRG	Year	All hospitals	w/o EMR	with EMR
	2006	1.2302	1.1778	1.3252
	2007	1.3201	1.2742	1.3724
	2008	1.2848	1.2129	1.3401
	2009	1.3164	1.2047	1.3539
	2010	1.3137	1.2189	1.3415
SURG DRG				
	2006	2.2868	2.2258	2.3868
	2007	2.2750	2.2484	2.3048
	2008	2.3972	2.3577	2.4232
	2009	2.4325	2.3780	2.4470
	2010	2.4562	2.4543	2.4566
MED DRG				
	2006	0.9749	0.9674	0.9885
	2007	0.9953	0.9882	1.0033
	2008	1.0356	1.0224	1.0458
	2009	1.0594	1.0427	1.0650
	2010	1.0583	1.0522	1.0601

Table 3.6: Coefficients of regression on Case-mix index

Pre-reform	CMI	SURG_CMI	MED_CMI
adopt	0.00614	0.0568**	-0.00314
	(0.00621)	(0.0228)	(0.00379)

Post-reform	CMI	SURG_CMI	MED_CMI
adopt	-0.00625	-0.0211*	-0.00540***
	(0.00459)	(0.0113)	(0.00169)

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 3.7: Coefficients of regression on the fraction

	Pre-reform	Post-reform
adopt	0.00197	-0.00987***
	(0.00265)	(-0.0019)
med	0.136***	-0.411***
	(0.00752)	(-0.00598)
adopt×med	-0.00480***	0.0115***
	(0.00159)	(-0.00181)

Note: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

APPENDIX A

Consider a linear regression model

$$Y = X\alpha + DX\beta + Z\gamma + u \tag{A.1}$$

where Y, X, Z are demeaned, N by 1 matrices with $E(XZ) \neq 0$; u is a classic error term with $E(u) = 0$ and $V(u) = \sigma_u^2$; and D is a $N \times N$ matrix representing a dummy variable. Therefore, DX is an interaction term and β captures the extra impact from X if the corresponding observation in D is equal to one. The model is well-defined in the sense that $E(u|X) = 0$, $E(u|D) = 0$, and $E(u|Z) = 0$. Assuming Z is not observable in the data, rewrite the model as follows

$$Y = X\alpha + DX\beta + \varepsilon \tag{A.2}$$

where $\varepsilon = Z\gamma + u$. As a result, the actual model suffers from endogeneity issue: $E(\varepsilon|X) \neq 0$ and $E(\varepsilon|DX) \neq 0$. This is a simplification of the model estimated in the paper.

Proposition 1 The Least Squares estimates for the real model has the following property: the bias for $\hat{\alpha}$ depends on the correlation between X and Z, as well as the sign of γ.

Proof. Consider $\hat{\alpha}$

$$\hat{\alpha} = (X'M_{DX}X)^{-1}X'M_{DX}Y \tag{A.3}$$

where

$$M_{DX} = I - DX[(DX)'DX]^{-1}(DX)' \tag{A.4}$$

is symmetric and idempotent. Replace Y with the true model:

$$\begin{aligned} \hat{\alpha} = & \quad (X'M_{DX}X)^{-1}X'M_{DX}(X\alpha + DX\beta + Z\gamma + u) \\ = & \quad \alpha + (X'M_{DX}X)^{-1}X'M_{DX}(DX)\beta + (X'M_{DX}X)^{-1}X'M_{DX}Z\gamma \\ & \quad +(X'M_{DX}X)^{-1}X'M_{DX}u \end{aligned} \tag{A.5}$$

Therefore,

$$E(\hat{\alpha}) = \alpha + \gamma E((X'M_{DX}X)^{-1}X'M_{DX}Z) \tag{A.6}$$

Note that $X'M_{DX}X > 0$ and therefore, the bias arises from $\gamma E(X'M_{DX}Z)$. Considering D is very similar to an identity matrix except that some of its diagonal elements are zero, $E(X'M_{DX}Z)$ only depends on the correlation between X and Z.

REFERENCES

Ackerberg, D. A. and G. Gowrisankaran (2006). Quantifying equilibrium network externalities in the ACH banking industry. *The RAND journal of economics*, **37**(3), pp. 738–761.

Adler-Milstein, J. and A. K. Jha (2014). No evidence found that hospitals are using new electronic health records to increase medicare reimbursements. *Health Affairs*, **33**(7), pp. 1271–1277.

Agha, L. (2011). The Effects of Health Information Technology on the Costs and Quality of Medical Care. Technical report, MIT Working Paper.

Aguirregabiria, V. and P. Mira (2003). Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models. *Econometrica*, **70**(4), pp. 1519–1543.

Aguirregabiria, V. and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica*, **75**(1), pp. 1–53.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Arcidiacono, P. and R. A. Miller (2011). CCP estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, **7**(6), pp. 1823–1868.

Bajari, P., C. L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica*, **75**(5), pp. 1331–1370.

Bowblis, J. R. and C. S. Brunt (2014). Medicare skilled nursing facility reimbursement and upcoding. *Health economics*, **23**(7), pp. 821–840.

Brunt, C. S. (2011). CPT fee differentials and visit upcoding under Medicare Part B. *Health economics*, **20**(7), pp. 831–841.

Collard-Wexler, A. (Forthcoming). Demand Fluctuations in the Ready-Mix Concrete Industry. *Econometrica*.

Coulam, R. F. and G. L. Gaumer (1992). Medicare's prospective payment system: a critical appraisal. *Health Care Financing Review*, **1991**(Suppl), p. 45.

Dafny, L. S. (2005). How Do Hospitals Respond to Price Changes? *The American Economic Review*, **95**(5), pp. 1525–1547.

Dranove, D., C. Forman, A. Goldfarb, and S. Greenstein (2012). The trillion dollar conundrum: Complementarities and health information technology. Technical report, National Bureau of Economic Research.

Ericson, R. and A. Pakes (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, **62**(1), pp. 53–82.

Gowrisankaran, G. and J. Stavins (2004). Network Externalities and Technology Adoption: Lessons from Electronic Payments. *RAND Journal of Economics*, pp. 260–276.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, **60**(3), pp. 497–529.

Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, **61**(2), pp. 265–289.

Jha, A., C. DesRoches, E. Campbell, K. Donelan, S. Rao, T. Ferris, A. Shields, S. Rosenbaum, and D. Blumenthal (2009). Use of electronic health records in US hospitals. *New England Journal of Medicine*, **360**(16), pp. 1628–1638.

Jha, A. K., D. Doolan, D. Grandt, T. Scott, and D. W. Bates (2008). The use of health information technology in seven nations. *International journal of medical informatics*, **77**(12), pp. 848–854.

Kellermann, A. and S. Jones (2013a). What it will take to achieve the as-yet-unfulfilled promises of health information technology. *Health Affairs*.

Kellermann, A. L. and S. S. Jones (2013b). What it will take to achieve the as-yet-unfulfilled promises of health information technology. *Health Affairs*, **32**(1), pp. 63–68.

Lee, R., J. McCullough, and R. Town (2012). The Impact of Health Information Technology on Hospital Productivity. Technical report.

Li, B. (2014a). Cracking the Codes: Do Electronic Medical Records. Technical report, Northwestern University Working Paper.

Li, B. (2014b). Cracking the codes: do electronic medical records facilitate hospital revenue enhancement.

Makuc, D., B. Haglund, D. Ingram, J. Kleinman, and J. Feldman (1991). Health service areas for the United States. *Vital and health statistics. Series 2, Data evaluation and methods research*, (112), p. 1.

Maskin, E. and J. Tirole (1988). A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pp. 571–599.

McCullough, J., M. Casey, I. Moscovice, and S. Prasad (2010). The effect of health information technology on quality in US hospitals. *Health Affairs*, **29**(4), pp. 647–654.

Miller, A. and C. Tucker (2011). Can health care information technology save babies? *Journal of Political Economy*, **119**(2), p. 289.

Miller, A. R. and C. Tucker (2014). Health information exchange, system size and information silos. *Journal of health economics*, **33**, pp. 28–42.

Orszag, P. (2008a). Evidence on the costs and benefits of health information technology. *Testimony before Congress*.

Orszag, P. R. (2008b). Evidence on the costs and benefits of health information technology. In *Testimony before Congress*, volume 24.

Pakes, A., M. Ostrovsky, and S. Berry (2007). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *The RAND Journal of Economics*, **38**(2), pp. 373–399.

Pesendorfer, M. and P. Schmidt-Dengler (2008). Asymptotic least squares estimators for dynamic games. *The Review of Economic Studies*, **75**(3), pp. 901–928.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, pp. 999–1033.

Ryan, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica*, **80**(3), pp. 1019–1061.

Silverman, E. and J. Skinner (2004). Medicare upcoding and hospital ownership. *Journal of health economics*, **23**(2), pp. 369–389.

Tucker, C. (2008). Identifying Formal and Informal Influence in Technology Adoption with Network Externalities. *Management Science*, **54**(12), pp. 2024–2038.

Wang, Y. (2012). *Cooperation and competition: the multilevel adoption of EMR in US hospitals*. Ph.D. thesis, Boston University.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Wynn, B. O., M. Beckett, L. H. Hilborne, M. Scott, B. Bahney, C. for Medicare & Medicaid Services, et al. (2007). *Evaluation of severity-adjusted DRG systems: interim report*. RAND.