Technological Diffusion Across Hospitals: The Case of a Revenue-Generating Practice

Adam Sacarny*

May 22, 2015

Abstract

Productivity-raising technologies tend to diffuse slowly, particularly in the health care sector. To understand how incentives drive adoption, I study a technology that generates revenue for hospitals: the practice of submitting detailed documentation about patients. After a 2008 reform, hospitals were able to raise their total Medicare revenue over 2% by always specifying a patient’s type of heart failure. I find that hospitals only captured around half of this revenue, indicating that large frictions impeded takeup. The key barrier is a principal-agent problem, since doctors supply the valuable information but are not paid for it. Exploiting the fact that many doctors practice at multiple hospitals, I find that four-fifths of the dispersion in adoption reflects differences in the ability of hospitals to extract documentation from physicians. Hospital adoption is also robustly correlated with the ability to generate survival for heart attack patients and the use of inexpensive survival-raising standards of care, suggesting that principal-agent problems drive disparities in quality more generally. These findings highlight the importance of agency conflicts in explaining variations in health care performance.

JEL Classification: D22; I1; O31; O33; L2

*Harvard University and Robert Wood Johnson Foundation, sacarny@rwj.harvard.edu. I am grateful to Amy Finkelstein, Michael Greenstone, Jon Gruber, and Paulo Somaini for their advice and guidance on this project. I thank Isaiah Andrews, Emily Baneman, David Chan, Manasi Deshpande, Amos Dodi, Kate Easterbrook, Ben Feigenberg, Eliza Forsythe, Paul Goldsmith-Pinkham, Sally Hudson, Greg Leiserson, Conrad Miller, David Molitor, Dianna Ng, Iuliana Pascu, Maxim Pinkovskiy, Maria Polyakova, Miikka Rokkanen, Annalisa Scognamiglio, Brad Shapiro, Henry Swift, Melanie Wasserman, Nils Wernerfelt, and participants in the MIT Public Finance lunch for their comments and suggestions. I would also like to thank Jean Roth for her assistance with all aspects of the Medicare data. I gratefully acknowledge funding from the National Institute on Aging grant T32-AG000186.
1 Introduction

Technology is usually believed to be a key driver of cross-country income disparities and economic growth. A classic finding of studies of technology is that new forms of production diffuse slowly and incompletely. For example, Griliches (1957) observed this pattern in the uptake of hybrid corn across states; more recent research has studied adoption patterns in agriculture in the developing world, manufacturing in advanced economies, management practices internationally, and a host of other examples (Conley and Udry, 2010; Foster and Rosenzweig, 1995; Collard-Wexler and Loecker, 2013; Bloom et al., 2012). Given the enormous productivity gains that result from many of these technologies, the nearly ubiquitous finding of delayed takeup is particularly vexing.

In this paper, I study a health care technology that raises revenue for the hospital: the detailed reporting of heart failure patients. A 2008 Medicare policy change created a financial incentive for hospitals to provide more detail about their patients in insurance reimbursement claims. Yet hospitals could only provide these details if they were documented by physicians. By tracking the diffusion of the reporting practice across hospitals, this study examines the role of financial incentives and agency conflicts in the adoption of new technologies.

These incentives are particularly important in the wake of the Affordable Care Act, which mandates that public insurers use their purchasing power to raise the productivity of health care providers.\(^1\) In designing new payment schemes, policymakers have focused on differences in the utilization of survival-raising processes of care, including checklists, hand-washing, and drugs like β-blockers. Disparities in the use of these practices are a leading explanation for health care productivity variations across providers and regions (Skinner and Staiger, 2009; Baicker and Chandra, 2004; Chandra et al., 2013). These processes of care require the coordination of hospitals and physicians, creating agency conflicts like those in the reporting of heart failure. While improved heart failure billing is a revenue-raising but not survival-raising technology, it is a clear test case of how financial incentives drive diffusion in the presence of agency frictions.

Hospitals have the option of listing heart failure on a reimbursement claim with detailed codes that describe the type of heart failure, or they may submit a vague code that provides little additional

\(^1\) The most prominent example of a policy resulting from that provision of the act is Medicare’s Value-Based Purchasing Program, which rewards hospitals that adopt evidence-based standards of care and perform well in surveys of patient satisfaction.
information about the condition. A 2008 reform\textsuperscript{2} changed the pricing function of Medicare to begin providing additional payments for the detailed codes. To capture this reward, hospitals needed to change how they reported their patients to Medicare. However, they could only make this change if doctors provided them with extra documentation about the heart failure to support it. The incentive for hospitals to report the information was large: this policy put over 2\% of hospital Medicare incomes on the line in 2009 – about $2 billion – though it did not affect the pay of physicians.

Figure 1 shows that the change in incentives triggered a rapid but incomplete response by hospitals: in just the weeks following the reform, hospitals started capturing 30\% of the revenue made available; by the end of 2011 they were capturing about 55\%. Presented inversely, in spite of the reform being announced earlier that year, 70\% of the extra heart failure revenue was not captured shortly after implementation and nearly half was still not being realized after several years.

I show that substantial hospital-level heterogeneity underlies the national takeup of detailed heart failure codes. Mirroring recent work that has demonstrated large differences in productivity across seemingly similar firms (Fox and Smeets, 2011; Doms and Bartelsman, 2000; Syverson, 2011), I find dispersion in the takeup of detailed billing codes across hospitals. This dispersion exists even after accounting for disparities in the types of patients that different hospitals treat. For example, 54\% of heart failure patients received a detailed code at the average hospital in 2010, and with the full set of patient controls the standard deviation of that share was 15 percentage points. A hospital two standard deviations below the mean provided detailed heart failure codes for 24\% of its heart failure patients, while a hospital two standard deviations above the mean did so for 84\% of its patients.

These findings suggest that hospitals were aware of the financial incentive to use the detailed codes, but that this awareness was tempered by significant frictions. I focus on frictions due to agency problems between a hospital and its doctors. Physicians are responsible for writing down the extra information about the heart failure, but Medicare does not pay physicians for the detailed codes or anything else that might be produced from the information.

\textsuperscript{2}All years are federal fiscal years unless otherwise noted. A federal fiscal year begins on October 1 of the previous calendar year, i.e. three months prior to the calendar year.
The principal-agent problem that this reform invokes is a classic one in economics – in other settings, it has been suggested as a driver, for example, of the failure of high quality management practices to diffuse across firms (Gibbons and Henderson, 2012). It plays a particular role in the American health care system because hospitals and physicians are frequently paid on independent bases. Moreover, with few exceptions, hospitals are not allowed to give a share of their Medicare payments to physicians as incentive pay. In spite of these restrictions, new policies to improve the quality of care have focused on the hospital’s payment alone.

The agency issues created by this reform arose from the bifurcated payment system. Hospitals – the principals – had large incentives to submit detailed codes about their patients, while physicians – the agents – had no direct incentive to provide the information. To resolve the principal-agent problem, hospitals would need to work with their doctors to better document their patients’ conditions, then translate this documentation into the newly valuable specific codes.

To study the role of these agency problems, I consider adoption rates that control for physician effects. Because doctors practice at multiple hospitals, it is possible to decompose the practice of detailed documentation into hospital- and physician-specific components. This decomposition is a novel application of a labor economics technique that has been frequently used in the context of workers and firms (see e.g. Abowd et al., 1999; Card et al., 2013) but has rarely been applied in studies of health.

Sweeping out the physician contribution removes dispersion in adoption due to hospitals having different kinds of doctors. This procedure addresses the concern that doctors who work at some hospitals may be more willing to provide the details than doctors who work at other hospitals. I show that dispersion is, if anything, slightly increased when the hospital component is isolated: the standard deviation of the share of patients who received detailed documentation across all hospitals rises from 0.15 percentage points with rich patient controls to 0.16 percentage points with patient and physician controls.\(^3\) The presence of residual variation means that even if facilities had the same doctors, some would be more capable of extracting specific documentation from their physicians than others. This result raises the possibility that institution-level principal-agent problems underlie some of the productivity differences that have been found among seemingly similar enterprises.

\(^{3}\)When there is negative assortative matching between hospitals and physicians, dispersion in adoption can rise when the physician component is removed.
I also consider the correlation between hospital adoption – with physician effects removed – and hospital characteristics like size, ownership, location, and productivity. The signs of these relationships are not \textit{ex ante} obvious, and they indicate which types of hospitals were most able to extract the codes from their doctors. The most robust finding of this analysis is that adoption was greater among hospitals that were higher quality by two measures: heart attack treatment productivity (the survival rate of heart attack patients after adjusting for spending on medical inputs and patient characteristics) and utilization of inexpensive, survival-raising processes of care (which includes administering aspirin after heart attacks and providing antibiotics before high-risk surgeries, among other evidence-based interventions). Under the view that extracting the revenue-generating codes from physicians makes a hospital revenue-productive, these results show that treatment and revenue productivity are positively correlated.

In an additional exercise, I look at correlates of hospital-level adoption but do not remove the component of takeup that is due to the hospital’s physicians. These results indicate which types of hospitals are the most “policy-elastic” with respect to financial incentives. Adoption is strongly correlated with hospital size, ownership, and productivity. Large and non-profit facilities were more likely to adopt, as were facilities that complied with consensus survival-raising standards of care. Hospital responses were also positively correlated with heart attack treatment productivity. This result touches on a key policy implication of this study, that financial incentives that push providers to raise treatment quality may be relatively ineffective on the low quality facilities most in need of improvement.

I contribute to the growing literature on productivity disparities and technological diffusion in three novel ways. First, by focusing on whether hospitals are able to modify their billing techniques to extract revenue, I isolate disparities in a context where it is uniquely plausible that none might exist. These disparities reflect differences in hospitals’ basic ability to respond to incentives. Second, using decomposition techniques that are normally associated with labor economics, I show that variations in adoption are largely driven by the ability of some hospitals to extract more high-revenue codes from their doctors than others – disparities persist when the physician component of adoption is removed. Lastly, I correlate the adoption of revenue-generating codes with the use of high quality standards of care in treatment to find that a common factor may drive both outcomes. Taken together, these findings hint that principal-agent problems may play a role in productivity
dispersion more generally – inside and outside the health care sector.

The paper proceeds as follows. Section 2 discusses the heart failure billing reform, the data I use to study it, and provides a simple analytical framework. Section 3 presents results on dispersion in hospital takeup, then shows how takeup relates to hospital characteristics and measures of treatment productivity. Section 4 provides a discussion of the results. Section 5 concludes.

2 Setting and Data

Heart failure (HF) is a syndrome defined as the inability of the heart’s pumping action to meet the body’s metabolic needs. It is uniquely prevalent and expensive among medical conditions. There are about 5 million active cases in the United States; about 500,000 cases are newly diagnosed each year. Medicare, the health insurance program that covers nearly all Americans age 65 and over, spends approximately 43% of its hospital and supplementary insurance dollars treating patients who suffer from HF (Linden and Adler-Milstein, 2008). Limiting to hospital expenditures, the program spends more on diagnosing and treating patients with HF than on patients with heart attacks. HF spending also outstrips spending on patients with all forms of cancer combined (Massie and Shah, 1997).

Medicare’s payment for heart failure is especially consequential for health expenditures and salient to hospital administrators, yet most economic literature on health care eschews studying HF in favor of less common conditions like heart attacks. The literature has focused on these conditions because they are thought to be sensitive to treatment quality and are well observed in most administrative data. Since this paper concerns how hospitals learn to improve their billing practices, not the effect of treatment on health, issues like endogenous selection of patients on unobserved determinants of survival are not the principal potential confounders. Rather, the great deal of revenue at stake for heart failure reimbursement makes it a condition that is well suited for this study’s aim of understanding how hospitals respond to coding incentives.

My analyses focus on the revenue generating practice of better documenting HF on hospital inpatient reimbursement claims to Medicare. The hospitals I study are paid through Medicare’s Acute Inpatient Prospective Payment System (IPPS), a $111 billion program that pays for most Medicare beneficiaries who are admitted as inpatients to most hospitals in the United States (MEDPAC,
As part of a 2008 overhaul of the IPPS – the most significant change to the program since its inception – the relative payment for vaguely documented and specifically documented HF was changed. This element of the reform made the documentation valuable and provided the financial incentive for the spread of the technology.

2.1 Payment Reform for Patient Documentation

The 2008 overhaul was a redesign of the IPPS risk-adjustment system, the process that adjusts payments to hospitals depending on the severity, or level of illness, of a patient. Medicare assigns a severity level to every potential condition a patient might have. A patient’s severity is the highest-severity condition listed on his hospital’s reimbursement claim. The reform created 3 levels of severity (low, medium, or high) where there had been 2 (low or high), shuffling the severity level of the many heart failure codes in the process.

By the eve of the reform, Medicare policymakers had come to believe that the risk-adjustment system had broken down, with nearly 80% of inpatients crowded into the high-severity category (GPO, 2007; Dafny, 2005 studies how hospitals exaggerate their reporting of patient severity due to incentives; Song et al., 2010 studies how reporting varies across regions). The reporting of HF had been a primary cause of the breakdown: there were many codes describing different types of HF, and all of them had been considered high-severity. Patients with HF accounted for about 25% of high-severity patients (or 20% of patients overall) in 2007.

Risk adjustment relies on detailed reporting of patients by providers, but according to the Centers for Medicare & Medicaid Services (CMS), the agency that administers Medicare, the overwhelmingly most common of the HF codes – 428.0, “congestive heart failure, unspecified” – was vague. Moreover, patients with this code did not have greater treatment costs than average (GPO, 2007). A set of heart failure codes that gave more information about the nature of the condition were found to predict treatment cost and, being specifically identified illnesses, were medically consistent with the agency’s definitions of medium and high severity. The vague code was moved to the low-severity list, but each of the detailed codes was put on either the medium- or the high-severity list. These codes and their severity classifications are listed in Table 1.

The detailed codes were exhaustive over the types of heart failure, so with the right documentation, a hospital could continue to raise nearly any HF patient to at least a medium level of severity.
following the reform. The specific HF codes indicate whether the systolic or diastolic part of the cardiac cycle is affected and, optionally, whether the condition is acute or chronic. Submitting them is a process that requires coordination between physicians and hospital staff. In this way it is similar to other technologies that have come into the focus of researchers and policymakers recently, including the use of β-blockers (an inexpensive class of drugs that have been shown to raise survival following a heart attack) in health care and the implementation of best managerial practices in firms.

For a hospital to legally submit a detailed code, a doctor must state the details about the HF in the patient’s medical chart. Figure 2 presents a flowchart of the organizational processes involved in the coding of patients. As the physician treats a patient, she writes information about diagnoses, tests, and treatments in the patient’s medical chart. When the patient is discharged, the physician summarizes the patient’s encounter, including the key medical diagnoses that were confirmed or ruled out during the stay. This discharge summary provides the primary evidence that the hospital’s health information staff (often called coders) use when processing the chart (Youngstrom, 2013). The staff can review the chart and send it back to the doctor with a request for more information – this process is called querying. Then, the staff must convert the descriptions of diagnoses into the proper numeric diagnosis codes, which becomes a part of the inpatient reimbursement claim (a concise description of the coding process can be found in O’Malley et al., 2005).

Both physicians and staff needed to revise old habits and learn new definitions; they also needed to work together to clarify ambiguous documentation. Coding staff might query a physician to specify which part of the cardiac cycle was affected by the HF, and other staff might review patient charts and instruct physicians on how to provide more detailed descriptions.

2.2 Revenue at Stake from Reform

Since HF was so common and the payment for having a medium- or high-severity patient was so much higher than the low-severity payment, hospitals had a clear incentive to use detailed codes whenever possible. Before the reform, the gain from these detailed codes relative to the vague code

---

4The chart is a file, physical or electronic, containing the patient’s test results, comments by providers of treatment, and ultimately a set of primary and secondary diagnoses. Its role is to provide a record of the patient’s stay for the purposes of treatment continuity and coordination, but the chart also serves as documentation supporting the hospital’s claims on payers like Medicare. CMS and its contractors frequently review charts to ensure that providers are not “upcoding”, or submitting high-paying codes that are not indicated by the documentation.
was zero because they were effectively identical in the Medicare payment calculation. Consistent
with these incentives, fewer than 15% of HF patients received a specific code in the year before the
reform.

Following the reform, the gain was always weakly positive and could be as high as tens of
thousands of dollars; the exact amount depended on the patient’s main diagnosis and whether the
patient had other medium- or high-severity conditions. For patients with other medium-severity
conditions, hospitals could gain revenue if they could find documentation of a high-severity form of
HF. For patients with other high-severity conditions, finding evidence of high-severity HF would not
change Medicare payments. However, using the detailed codes was still beneficial to the hospital
because it would help to keep payments from being reduced if the claim were audited and the other
high-severity conditions were found to be poorly supported.

In 2009, the average gain per HF patient from using a detailed HF code instead of a vague one
was $227 if the code indicated chronic HF (a medium-severity condition) and $2,143 if it indicated
acute HF (a high-severity condition). As a point of comparison, Medicare paid hospitals about
$9,700 for the average patient and $10,400 for the average HF patient in 2009. The evolution of the
gain to specific coding is shown in Figure 3 and the corresponding takeup in the use of these codes
is shown in Figure 4.

For each hospital, the gain to taking up the revenue-raising technology – the money put on the
table by the reform – depended on its patient mix. Hospitals with more HF patients, and more
acute (high-severity) HF patients, had more to gain from adopting specific HF coding. To get a
sense of how this gain varied across hospitals, I predict each hospital’s \textit{ex ante} revenue put at stake
by the reform. This prediction takes the hospital’s 2007 HF patients, probabilistically fills in the
detailed HF codes the patients would have received under full adoption of the coding technology,
and determines the ensuing gain in payment from these codes by processing the patient under the
new payment rules. Heart failure codes are predicted using the relationship between coding and
patient characteristics in hospitals that were relatively specific coders in 2010.\footnote{These averages include the patients for whom the detailed codes do not raise payments because, for example, they already had another medium- or high-severity condition. To determine how a hospital would have been paid had it coded HF differently, I use a computer program called a grouper that translates an inpatient claim into its Medicare payment diagnosis-related group (DRG). The gains to specific HF codes were calculating by reprocessing all HF patients, replacing the observed HF codes with only vague, only chronic, and only acute HF codes.}

\footnote{This predictor uses HF patients at hospitals that were relatively detailed coders in 2010 – hospitals that gave at least 85% of their HF patients a detailed code. The sample includes 90,653 patients and 171 hospitals. I regress}
Figures 5 and 6 show the high level of and variation in *ex ante* revenue put at stake by the reform across hospitals; the average hospital would have expected to gain $1,007 per HF patient (or, spreading this gain across all admissions, $268 per patient) in 2009 by giving all of its HF patients specific HF codes rather than vague ones. The standard deviation of the revenue at stake per HF patient was $230 (the standard deviation of the gain spread over all patients was $76). To provide a sense of scale, one can consider these amounts relative to hospital operating margins. The 2010 Medicare inpatient margin, which equals hospitals’ aggregate inpatient Medicare revenues less costs, divided by revenues, was -1.7% (MEDPAC, 2012a). This negative operating margin has been cited by the American Hospital Association as evidence that Medicare does not pay hospitals adequately (American Hospital Association, 2005). The gains from detailed coding for HF were even larger than this margin: pricing the pre-reform patients under the 2009 rules shows that hospitals could have expected to raise their Medicare revenues by 2.9% by giving all of their HF patients specific HF codes.

2.3 Analytical Approach

The basic framework for analyzing takeup of the technology views the decision to use a specific HF code $code \in \{0, 1\}$ as a function of the propensity of the hospital and the doctor to favor putting down the code or documentation thereof. I let hospitals be indexed by $h$, doctors by $d$, and patients by $p$. Under the assumption of additive separability of the hospital and the doctor’s effects on the coding probability, hospitals can be represented by a hospital type $\alpha_h$ and doctors by a doctor type $\alpha_d$. Patient observables are $X_p$ and the remaining heterogeneity, which accounts for unobserved determinants of coding behavior, is $\epsilon_{ph}$:

$$code_{ph} = \alpha_h + \alpha_d + X_p \beta + \epsilon_{ph}$$

(1)

The hospital’s type can be thought of as its underlying propensity to extract specific HF codes independently of the types of physicians who practice at the hospital. The doctor type reflects that whether the patient was coded as having high-severity HF on well-measured patient attributes (indicators for: age, race, sex, month of admission, whether admitted through the emergency department, 19 chronic conditions, and 25 major diagnostic categories classifying the underlying cause of admission). I use this regression to fit the probability that patients in 2007 would have received a medium- or high-severity HF code, then re-price these patients under the 2009 post-reform pricing rules. The result of this procedure is an *ex ante* expected gain to using the detailed codes, which I aggregate to the hospital level.
some physicians are more or less prone to document the kind of HF that their patients have due to their own practice styles and the incentives of the physician payment system. In this framework, doctors carry their types across hospitals. Finally, the patient component accounts for observed differences that, in a way that is common across facilities, affect the cost of providing a specific code.

The dispersion of the hospital types is of direct interest, and is the first focus of the empirical analysis. A wide literature has documented persistent productivity differentials in the manufacturing sector (see Syverson, 2011 for a review), and work is ongoing to develop documentation of similar facts in the service and health care sectors (Fox and Smeets, 2011; Chandra et al., 2013). In this framework, a hospital’s type can be thought of as its revenue productivity – its residual ability to extract revenue from Medicare after accounting for the observable inputs to the coding production process, like patient and doctor types. Dispersion in hospital types is therefore a form of productivity dispersion.

What might drive this dispersion? Recall that hospitals were constrained from directly incentivizing their doctors to provide the additional documentation needed to submit a specific HF code. When a doctor moves from a low-type hospital to a high-type hospital, her HF patients become more likely to have a detailed code, regardless of the doctor’s type. One perspective is that this difference is due to the high-type hospital better solving the principal-agent problem. The variation in hospital types can reflect variation in whether hospitals can bring their doctors’ behaviors in line with the hospital’s incentives.

The second element of the empirical analysis focuses on describing the kinds of hospitals that are most effective at responding to the incentives for detailed coding. These analyses look at the relationships between hospital types and characteristics of the hospital. The first set of characteristics, called $C_h$, comprises the hospital’s size (defined as the number of beds in the facility), ownership (non-profit, for-profit, or government-run), location (whether the hospital is in a large urban area, other urban area, or rural area), teaching status (whether it has residents), and ex-ante per-patient revenue put at stake by the reform. The second set, called $Z_h$, includes measures of the hospital’s productivity – the amount of survival the hospital can generate for a fixed amount of inputs.
In the key hospital-level analysis, I regress the hospital type on these two sets of characteristics:

$$
\alpha_h = \gamma + C_h \rho + Z_h \theta + \eta_h
$$

The signs of the elements of $\rho$ and $\theta$ are not obvious, both because the causal relationships between hospital characteristics and the takeup of revenue-generating technology are not well known and because other, unobserved factors may be correlated with $C_h$ and $Z_h$ and drive takeup. I discuss these potential relationships and estimate this equation in Section 3.

### 2.4 Data

I study the impact of the IPPS reform on the diffusion of the revenue generating practice using a dataset of all inpatient hospitalizations for Medicare beneficiaries. My data is primarily drawn from the MEDPAR Research Information File (RIF), a 100% sample of all inpatient stays by Medicare beneficiaries with hospital care coverage through the government-run fee-for-service system. This file is essentially a copy of all the reimbursement claims that hospitals sent Medicare. For 92% of these stays, I can identify the physician who was primarily in charge of taking care of the patient in the hospital and thus most responsible for the final diagnoses that were coded and submitted on the hospital's claim. Since physicians are paid for each procedure they perform, for these stays I can also identify echocardiograms and other heart tests.

I use these data to construct an analysis sample of hospitals' claims to Medicare for their HF patients. Starting with all patients in 2010, I eliminate those who lacked full Medicare coverage at any point during their hospital stay, were covered by a private plan, or were under age 65. To focus on hospitals that were subject to the reform, I include only inpatient acute care facilities that are paid according to the IPPS. As a result, I drop the approximately 3% of stays that occur at Critical

---

7I use the attending physician identifier from the Medicare Inpatient RIF. To ensure that only valid individual physicians are included, I drop physician identifiers that could not be found in the AMA Masterfile, a census of all physicians, which accounts for most of the stays for which the physician was not observed.

The small literature on identifying the attending physician in Medicare claims has suggested looking at physician claims (found in the Medicare Carrier RIF) and choosing the physician who bills Medicare for the most evaluation and management services, rather than the physician indicated by the hospital on its inpatient claim (Trude, 1992; Trude et al., 1993; Virnig, 2012). There are two advantages to using the hospital’s report, however. First, the hospital’s report of the attending physician may more accurately reflect the physician with whom the facility was communicating to determine the patient’s diagnosis codes. The literature on identifying the physician is more concerned with the most medically responsible physician, not the one most responsible for billing and coding. Second, I only observe physician claims for a 20% random sample of patients, dramatically restricting the set of patients for whom I observe the physician when using the physician claim method.
Access Hospitals (these hospitals number about 1,300 but are very small and have opted to be paid on a different basis) and 2% of stays at Maryland hospitals (which are exempt from the IPPS).

I then limit the sample to heart failure patients, which are identified as those with a principal or secondary diagnosis ICD-9 code of 428.x, 398.91, 402.x1, 404.x1, or 404.x3.

This analysis sample is described in Table 2 and includes 1.9 million HF patients. There are about 3,400 distinct hospitals in the full analysis sample, which is the set of claims that is used to estimate hospital types when physician effects are not being swept out. Most hospitals see a large number of HF patients in a given year: the average treats 553 of them, and even the 10th percentile includes 52 of them.

By using the 92% of patients for whom I observe the physician, I am able to identify hospital effects after controlling for the doctor. Hospital and physician types are only separately identified within a “mobility group” – the set of hospitals and physicians that can be connected, in graph theory terms, by physicians who work at multiple facilities (this concept is explained in greater detail in Section 3.2). I call the mobility sample the set of patient claims that occur within the largest mobility group of hospitals and physicians. There are about 2,900 hospitals and 135,000 doctors in the sample. The average mobility sample hospital sees 582 HF patients in 2010 and its HF patients are treated by 58 distinct doctors. At the average hospital, 20 of these doctors are mobile, which means that they are observed treating at least one HF patient at another hospital. Mobile doctors are crucial for my analyses because their behavior separately identifies the hospital and doctor types.

In this sample, the average doctor sees 12 HF patients in a given year and works at 1.23 distinct hospitals. About 19% of doctors are mobile. Table 3 provides additional information about the doctors by mobility status. The average mobile physician treats about twice as many patients as a non-mobile physician. Information on physician specialty, demographics, training, and experience comes from the AMA Masterfile; specialties are grouped according to the Dartmouth Atlas definitions. Mobile physicians are 11pp less likely to be surgeons and are correspondingly more likely to be primary physicians like internists or medical specialists like cardiologists. Mobile physicians also have about 8 months more training – but about 8 months less experience practicing

\[8\text{See Table 2 of the document found at http://www.dartmouthatlas.org/downloads/methods/research\_methods.pdf}\]
since completing training – than their non-mobile counterparts.

2.5 Costs of Takeup

Figure 1 shows that the large amount of revenue at stake for specific coding induced an almost instantaneous partial takeup of the coding. Over the following years the takeup continued, though it remained far from 100% even by the end of 2011. The finding of incomplete takeup raises the question of what costs must be incurred by the hospital to adopt the technology.

One possibility is that taking up the reform requires medical testing of HF patients to confirm the details of their conditions. For example, the gold standard for confirming whether there is systolic or diastolic dysfunction – the minimum amount of information needed to use a specific code – is an echocardiogram, a non-invasive diagnostic test. Some observers proposed that the reform put pressure on physicians to perform echocardiograms that they had not considered medically necessary (Leppert, 2012). If these concerns were true, one could interpret the reform as encouraging higher intensity medicine – and the costs and principal-agent frictions as the refusal of doctors and hospital staff to order tests that they had not already thought necessary.

Contrary to this story, both the time series evidence and the official coding guidelines show that whatever were the costs of more detailed HF coding, they were not realized through changes in real medical treatment. Figure 7 shows that the enormous increase in the capture of HF coding revenue was not matched by any perceptible change in heart testing as measured by the share of all patients receiving an echocardiogram. This finding is sensible considering that with enough information to diagnose and submit a vague HF code, it is almost always possible to provide enough additional documentation to legally submit a specific HF code: a patient’s medical history and symptoms are predictive of the type of HF, and the official coding guidelines state that “if a diagnosis documented at the time of discharge is qualified as ‘probable,’ ‘suspected,’ ‘likely,’ ‘questionable,’ ‘possible,’ or ‘rule out,’ the condition should be coded as if it existed or was established” (Prophet, 2000). Thus these codes require only suggestive evidence, not the certainty of an echocardiogram.

A key source of takeup frictions comes from a principal-agent problem that pitted a hospital interest in detailed documentation against physicians who had little to gain financially from providing the information. Although this documentation may seem nearly costless to produce, physicians face many competing demands on their time when they edit medical charts. HF is often just one
condition among many that are relevant to the patient’s treatment. For example, a doctor’s first-order concern may be documenting aspects of the patient that are crucial for proper post-acute care, making documentation that matters solely for the hospital’s billing a secondary issue.

Taking up the revenue-generating technology required hospitals to pay a variety of fixed and variable costs that were unrelated to patient treatment but could influence physicians’ documentation styles. Examples of these costs include training hospital staff to prompt doctors for more information when a patient’s chart lacks details and purchasing health information technology that prompts staff to look for and query doctors about high-value codes. Hospitals also could expend resources creating ordeals for physicians who fail to provide detailed documentation. The view that physician habits are expensive for the hospital to change matches accounts of quality improvement efforts that sought to make reluctant physicians prescribe evidence-based medicines, wash their hands, and perform other tasks to improve mortality and morbidity (Voss and Widmer, 1997; Stafford and Radley, 2003; Pittet et al., 1999).

3 Hospital Adoption

Incentive misalignments owing to principal-agent problems have been proposed as impediments to the adoption of new technology and to making organizational change more generally. One notable example of this view is found in Gibbons and Henderson (2012), who adapt a typology of managerial pathologies, focusing in particular on the many failures of organizations to take up practices that were widely known to be beneficial. These failures, they argue, are consistent with poor implementation: managers “know they’re behind, they know what to do, and they’re trying hard to do it, but they nonetheless cannot get the organization to get it done.”

Implementation difficulties are particularly acute in the health care setting because facilities (in this context, the principals) and physicians (the agents) tend to be paid separately and on different bases. In the case of heart failure, physician payments from Medicare do not depend on whether a reimbursement claim uses vague or detailed diagnosis codes because physicians are paid for each procedure they perform. Though hospitals might want to encourage detailed coding by paying doctors for it, doing so would likely run afoul of federal laws that prohibit directly incentivizing physicians by basing their payments on the hospital’s payment (see HHS, 1999; this practice is
commonly known as gainsharing). In a sense, the principal-agent problem in patient documentation is literally written into the law.

The adoption of the coding technology was incomplete at the national level, but the national time series masks enormous heterogeneity at the level of the hospital. In this section, I construct and present measures of adoption of detailed coding across hospitals. These measures have wide dispersion, with some hospitals almost never using specific codes and other hospitals almost always using them. A perhaps natural view is that some health care providers are uniquely unable or unwilling to respond to incentives. Yet dispersion alone is not enough to make health care exceptional on the dimension of technology adoption – this finding is nearly universal in cases of new technology, and persistent differences in productivity have been found in nearly every sector in which they have been studied.

I present a novel analysis of the role that physicians played in the adoption of the revenue generating practice. I decompose the hospital’s average coding into the component that is due to the facility and the component that is due to its doctors. The notion of outcomes being due to a hospital and doctor component follows a commonly used econometric model of wages that decomposes them into firm and worker effects (see e.g. Abowd et al., 1999 and more recently Card et al., 2013, which study wages in France and Germany, respectively).

This section undertakes two key analyses. First, it shows that dispersion in the adoption of detailed HF coding persists even among observably similar hospitals. The dispersion result is robust to sweeping out the physician component of coding – even if hospitals had the same doctors, there would still be coding disparities. Equivalently, the probability a HF patient treated by a particular doctor gets a specific code is significantly greater at some hospitals relative to others.

Second, it explores the relationship between adoption and hospital characteristics like size, ownership, and quality. The signs of these relationships are not \textit{ex ante} obvious, but they speak to several important and open questions in health economics. Though these results are descriptive, not causal, they are useful policy inputs: they can be interpreted as indications of which providers are most elastic to incentives for revenue generating technologies.
3.1 Econometric Specification

The key analyses of this section describe the distribution of the adoption of the coding technology with two-step methods. The first step extracts a measure of adoption at the hospital level, which is the hospital effect given in equation 1. This fixed effect is the probability that a HF patient in the hospital receives a detailed HF code, after adjusting for patient observables and doctor effects. In the second step, I analyze the distribution of the fixed effects by calculating their variance (to look for variations among seemingly similar enterprises) and by regressing them on hospital characteristics and productivity (to see which facilities are most likely to adopt).

3.1.1 First Step: Estimating Hospital Fixed Effects

In the first step, I run the regression given in equation 1. I consider versions of this regression with patient controls of varying degrees of richness, and run these regressions both with and without physician fixed effects. I then extract estimates of the hospital fixed effects \( \hat{\alpha}_h \). These estimates equal the share of HF patients at the hospital who received a specific code \( \text{code}_h \) less the contribution of the hospital’s average patient \( (\bar{X}_h, \hat{\beta}) \) and the patient-weighted average physician effect \( \left( \frac{1}{N_h} \sum_{p \in P_h} \hat{\alpha}_{d(p)} \right) \), where \( N_h \) is the number of HF patients at the hospital, \( P_h \) indexes the patients, and \( d(p) \) indicates the doctor that attended to patient \( p \):

\[
\hat{\alpha}_h = \text{code}_h - \bar{X}_h \hat{\beta} - \frac{1}{N_h} \sum_{p \in P_h} \hat{\alpha}_{d(p)}
\]

In the simplest specification, which includes no patient controls nor physician fixed effects, the estimates of the hospital fixed effects \( \hat{\alpha}_h \) become the shares of HF patients in hospital \( h \) who receive a specific HF code:

\[
\hat{\alpha}_{h \text{ simple}} = \text{code}_h
\]  

(3)

There are two caveats to using this measure, both of which can be seen by taking the difference between \( \hat{\alpha}_{h \text{ simple}} \) and \( \hat{\alpha}_h \):

\[
\hat{\alpha}_h - \hat{\alpha}_{h \text{ simple}} = \bar{X}_h \hat{\beta} + \frac{1}{N_h} \sum_{p \in P_h} \hat{\alpha}_{d(p)}
\]
One is that heterogeneity in $\hat{\alpha}_{h}^{\text{simple}}$ may be due to patient-level factors $\bar{X}_h/\hat{\beta}$ that have been shifted to the error term of the simple measure. For example, dispersion in coding could reflect that some hospitals have patients who are difficult to code. The specifications with rich sets of patient observables account for this concern. When patient-level factors are included, the use of hospital (and potentially physician) fixed effects means that the coefficients on patient characteristics are estimated from the within-hospital (and potentially within-physician) relationships between these characteristics and coding.

The second caveat is that dispersion could also reflect the role of physicians in coding, $\frac{1}{N_h} \sum_{p \in P_h} \hat{\alpha}_d(p)$ – some hospitals may have doctors who are particularly willing or unwilling to provide detailed documentation of their patients. Whether the physician component should be removed depends on the analysis, since the physician’s actions inside the hospital are an endogenous component of the hospital’s response to the reform. For example, hospitals with much to gain from the reform may be more likely to teach their physicians how to recognize the signs and symptoms of HF. These physicians would then be more likely to document specific HF in any hospital. Controlling for the physician effects would sweep out this improvement. Still, the extent to which the response to the reform is driven by changes in hospital behavior above and beyond the actions of its physicians is of interest in identifying principal-agent problems.

3.1.2 Second Step: Describing the Distribution of the Hospital Fixed Effects

This section explains the analyses of the $\hat{\alpha}_h$ and how they account for estimation error due to sampling variance.

**Dispersion among Similar Hospitals** The first key analysis of this paper studies the dispersion of the hospital fixed effects. However, the objects $\hat{\alpha}_h$ are noisy – though unbiased – estimates of $\alpha_h$, meaning that their dispersion will be greater than the true dispersion of $\alpha_h$. This noise comes from small samples at the hospital level (some hospitals treat few HF patients) and imprecision in the estimates of the other coefficients in the model. When the specification lacks physician fixed effects, the other coefficients in the model are at the patient level, and are estimated from millions of observations. These coefficients are estimated quite precisely, reducing the role for this noise.

When the specification includes physician fixed effects, the imprecision of the hospital effect
reflects imprecision in the estimates of the physician effects. In a simple specification with no patient-level characteristics, the hospital effects are identified only by patients who were treated by mobile doctors, and one component of the measurement error in the hospital effect is an average of the measurement error of those physicians’ effects. As these coefficients become estimated more precisely, for example as the number of patients treated by the mobile doctors rises, the estimation error falls.

Estimates of the variance of $\alpha_h$ must account for measurement error in order to avoid overstating dispersion. To produce these estimates, I adopt part of the Empirical Bayes procedure described in Appendix C of Chandra et al. (2013). This procedure uses the diagonals of the variance-covariance matrix from the first-step regression as estimates of the variance of the hospital fixed effect measurement error. I generate a consistent estimate of the variance of $\alpha_h$ by taking the variance of $\hat{\alpha}_h$ and subtracting the average squared standard error of the hospital fixed effects (i.e. the average value of the diagonals of the variance-covariance matrix).

**Describing the Adopters** The other key analysis of this section describes the adopters by placing the hospital fixed effect estimates on the left-hand side of regressions of the form of equation 2. The measurement error in the $\hat{\alpha}_h$ therefore moves into the error term where its primary effect is to reduce the precision of the estimates of the coefficients $\rho$ and $\theta$. Since the measurement error is due to sampling variance in the first step, it is not correlated with the characteristics and productivity measures that are found on the right-hand side of the key regressions, and it does not bias the estimates of $\rho$ or $\theta$.

### 3.2 Separate Identification of Hospital and Physician

The health care context is unique because it allows the separate identification of the contribution of the principal and the contribution of the agent to takeup – a decomposition that cannot be performed when agents are observed under just one principal. The key insight behind the decomposition in the heart failure setting is that physicians are frequently observed treating patients at multiple hospitals, since doctors may have admitting privileges at several facilities. When the same physician practices in two hospitals, her propensity to provide detailed documentation at each facility identifies the hospital effects relative to each other. Likewise, when two physicians practice at the same hospital,
their outcomes at that hospital identify the physician effects relative to each other.

The physician fixed effects, when they are included in the first step, sweep out the component of the hospital’s coding that is due to the behavior of its doctors. The hospital and physician effects can be separately identified within a mobility group – the set of doctors and hospitals that are said to be “connected” to each other. Consider the graph of doctors and hospitals, in which each doctor and hospital is represented by a point (called a node in graph theory). In the graph, a doctor and hospital have a line (called an edge) drawn between their nodes if the doctor treats a patient at that hospital. Two hospitals or doctors are connected if there exists any unbroken sequence of lines (called a path) going from one to the other in the graph. A mobility group starts with a doctor or hospital and includes all other doctors and hospitals that are connected to her or it. In the graph of doctors and hospitals, a mobility group is called a maximal connected subgraph. Among the 3,414 hospitals in the analysis sample, the largest mobility group contains 2,868 hospitals.

The econometric model of the first step follows from certain identification assumptions. The key assumption is that the probability that a patient receives a specific code must approximate a linear probability model with additive effects from the patient, hospital, and doctor such that:

\[ \mathbb{E}[\text{code}_{ph}] = \alpha_h + \alpha_d + X_p\beta \]

Though the idea the three levels are linear and additively separable is only an approximation, the additivity assumption can be tested by estimating a match effects model (Card et al., 2013). This model replaces the hospital and physician fixed effects with a set of effects at the hospital-physician level (i.e. \( \alpha_{hd} \)), allowing any arbitrary relationship between hospital and physician types. The match effects model improves the explanatory power of the model minimally, suggesting that additivity is not a restrictive assumption in this context.\(^9\)

One implication of the conditional expectation equation is that patients do not select hospitals or doctors on the basis of unobserved costs of coding. If this were the case, for example, the fixed effect of a hospital with unobservably more costly patients would be estimated with negative bias. I test this assumption by including increasingly rich sets of patient characteristics as controls. The key

\(^9\)Specifically, the adjusted \( R^2 \) of the first-step regression with hospital fixed effects, physician fixed effects, and the full set of patient controls is 0.369, while the adjusted \( R^2 \) of the same regression with the two sets of fixed effects replaced by one level of hospital-physician match effects is 0.372.
results on the characteristics of the adopters and the dispersion in adoption are somewhat sensitive to these controls. Specifically, the significant coefficients in the regressions of adoption on hospital covariates tend to attenuate by at most one-third due to the inclusion of rich patient characteristics observable in the patient’s hospital billing claim, but they are not further reduced by including controls for patient histories of chronic illnesses (the controls are described in section 3.4). These coefficients remain highly significant even though they attenuate. Likewise, the standard deviation of adoption is reduced by about one-fourth from the patient controls, and again the reduction is entirely due to characteristics in the billing claim.

It is perhaps unsurprising that patient characteristics influence the hospital’s use of the codes. The fact that adding patient illness histories as an additional set of controls does not further affect dispersion in adoption suggests that the key factors are attributes of the patient’s admission. The identifying econometric assumption is that unobserved characteristics are not playing a role in coding, and the information observable about the admission is quite detailed in the claims data that I use. While the great majority of disparities in adoption across hospitals cannot be attributed to anything observable about the patient, I present all results in this study under three patient-level specifications to be clear about this potential source of endogeneity.

A related identification requirement is that the assignment of doctors to hospitals must not reflect match-specific synergies in the coding outcome. Though there may be an unobserved component of coding that is due to the quality of the match, the matching of doctors and hospitals must not systematically depend on this component. For example, a hospital might demand less specificity in HF coding from physicians who were friendly with its owners. These physicians would have negative match effects with that hospital. If they tended to practice at the hospital, the negative match effects would load onto the hospital effect, biasing it downward. The role of match-specific synergies is also bounded by the match effects model described in footnote 9 – the low explanatory improvement of that model indicates that the size of these synergies must be small, limiting the scope for endogeneity from this source.

3.3 Hospital Characteristics

Table 4 shows summary statistics for the cross section of hospitals that I include in the dispersion and characteristics of adopters analyses. This cross section consists of 2,411 hospitals, and includes
any facility with a heart failure coding score and which had complete information on all baseline characteristics, standards of care, and productivity.

The rows of the table comprise the key hospital characteristics and productivity measures that are used in the analyses. Hospital size (beds) and ownership are taken from the Medicare Provider of Services file. Ownership may be non-profit (about two-thirds of hospitals), for-profit (one-sixth), and government-run (one-sixth). Hospital location and teaching status are taken from the 2010 Medicare IPPS Impact file. The location definition is the one used by Medicare: a large urban area is any Metropolitan Statistical Area (MSA) with a population of at least 1 million, an other urban area is any other MSA, and the rest of the country is considered rural. The hospitals in this sample are found in all three areas, though the number of rural hospitals is reduced because many were classified as critical-access facilities, which were exempt from this reform. Teaching hospitals, which comprise just over one-third of facilities, are defined as those with a resident-to-bed ratio greater than zero.

I define the \textit{ex ante} revenue at stake as the expected value of giving all of the hospital’s pre-reform (2007) HF patients a specific code according to post-reform (2009) reimbursement rules. The revenue at stake is scaled by the total number of patients at the hospital, making it the per-patient expected gain from fully taking up the reform. Since most patients were coded vaguely in 2007, this variable is constructed by filling in each 2007 patient’s specific HF code using the relationship between well-observed patient characteristics and specific HF codes at hospitals that were excellent at coding in 2010 (see footnote 6 for more information). To improve precision and reduce the leverage of outliers, hospitals with fewer than 50 HF patients in 2007 as well as those with an outlying top or bottom 1% of revenue on the table per patient were culled from this measure.

Heart attack treatment productivity is constructed using the sample and methods of Chandra et al. (2013). A hospital’s treatment productivity is the average log-survival of heart attack patients treated at the hospital in 2000-2006, after controlling for the inputs used to treat the patient and a rich set of patient observables. Raising hospital productivity by 10% means that, at the same level of inputs, the hospital is able to produce 10% more survival-days for its patients. Productivity is adjusted to account for measurement error using an Empirical Bayes shrinkage procedure described in more detail in Appendix C of Chandra et al., 2013.

The standards of care measures were collected by CMS under its Hospital Compare program.
They indicate the shares of times that standards of care were followed for heart attack, heart failure, pneumonia, and high-risk surgery patients in 2006. These standards of care are inexpensive, evidence-based treatments that have been shown to improve patient outcomes. When productivity is defined as the amount of survival a hospital can generate for a fixed set of inputs, these scores measure the takeup of productivity-raising technologies. They notably include β-blockers, a class of inexpensive drugs that dramatically improve survival following heart attacks and that has been the subject of several economic studies (see e.g. Skinner and Staiger, 2009, 2007).

3.4 Dispersion

I find dispersion in adoption with and without rich patient and physician controls. To provide a sense of the time series of adoption, Figure 8 shows the distribution of raw $\hat{\alpha}_h^{\text{simple}}$, the share of HF patients at hospital $h$ who received a detailed HF code, in each year from 2003 to 2010. Takeup across hospitals moved rapidly after the reform. By 2010, the median hospital used specific codes 55% of the time. Figure 9 shows the full distribution of $\hat{\alpha}_h^{\text{simple}}$ in 2010, the analysis sample year. There was great variation in takeup across hospitals even in the third year following the reform. In particular, there was a substantial mass of hospitals using detailed codes less than 20% of the time, and a nontrivial number of hospitals that almost never used them. These figures plot the $\hat{\alpha}_h^{\text{simple}}$ with no adjustment for measurement error, but they exclude hospitals with fewer than 50 HF patients to limit the scope for measurement error to drive dispersion. All results shown in tables make adjustments for excess dispersion due to sampling variance, however.

Table 5 shows the standard deviation of adoption among observably similar hospitals. I divide the space of hospitals into 7 mutually exclusive and exhaustive groups on the basis of characteristics that have been the focus of literature on hospital quality. The table includes three sets of patient controls in the first step, which is where the hospital effects are extracted. In the left three columns, each patient control specification is presented without first-step physician effects; in these results, the hospital effects include the component of coding that is due to the physicians. The right three columns show the hospital effects adjusted for excess dispersion due to sampling variance.

---

The processes of care included in the measures were chosen by CMS based on medical evidence. The heart attack measure includes prescription rates of β-blockers and aspirin for appropriate patients as well as 5 other processes of care. The heart failure measure includes an evaluation of left ventricular systolic function (a key input to determining the part of the cardiac cycle that is weakened) and 3 other processes of care. The pneumonia measure includes prompt prescription of antibiotics and 6 other processes of care, and the surgery measure includes antibiotics and 2 other processes of care.
columns add first-step physician effects, which sweeps out the physician component.

The first patient control specification, presented in columns (1) and (4), includes no patient-level controls at all. The second, presented in columns (2) and (5), includes observables about the patient’s hospital admission found in the hospital’s billing claim: age, race, and sex interactions; whether the patient was admitted through the emergency department; and 179 categories for the patient’s primary diagnosis. The third set of patient controls, shown in columns (3) and (6), augments the second set to also include indicators for whether the patient had any of 19 chronic conditions.

The result of Table 5 is that dispersion shrinks somewhat as rich controls about the patient’s hospital admission are added in the first step, though controlling for patient illness histories has little effect. Moreover, the addition of physician effects does not systematically reduce these variations, and it even raises dispersion slightly in the full cross-section of hospitals.

Among all hospitals, the standard deviation of the coding scores with no controls is 0.20, meaning that a hospital with one standard deviation greater adoption gives 20pp more of its HF patients a specific HF code. This measure does not account for differences in patient or doctor mix across hospitals. With all patient controls included, the standard deviation falls to 0.15. This dispersion is the standard deviation across hospitals of the probability a HF patient gets a specific code, holding fixed the patient’s characteristics. It calculates adoption across hospitals after removing the component that can be explained by within-hospital relationships between patient observables and coding. Further adding physician fixed effects raises the standard deviation slightly to 0.16. This result is the dispersion across hospitals in the probability a specific code is used, given a HF patient with a fixed set of characteristics and a fixed physician. With these controls, a hospital with one standard deviation greater adoption is 16pp more likely to give a patient a specific code.

Within key groups of hospitals, dispersion tends to decline with the inclusion of patient characteristics in the first step; the additional inclusion of physician fixed effects may raise or reduce dispersion within these groups. Large, urban, non-profit teaching hospitals, for example, have a standard deviation in coding rates of 0.17 without any first-step controls, 0.13 with patient controls, and 0.12 with patient and physician controls. Likewise, the standard deviation of coding rates among non-urban non-profit teaching hospitals falls from 0.17 with no controls to 0.13 with patient controls, but rises to 0.19 when physician controls are further added. These patterns are replicated
in the other groups of hospitals: dispersion declines by 5-6pp with the inclusion of patient characteristics, but may decline (up to 1pp) or rise (up to 3pp) with the additional inclusion of physician effects.

While it may seem counterintuitive that disparities in adoption sometimes increase with the addition of physician controls, this finding is possible if high type hospitals tend to match with low type physicians. When physician controls are omitted, the hospital’s adoption includes both the facility component and an average physician component. Adding the physician controls removes the average physician component. When dispersion in adoption rises when these controls are added, it indicates that the average physician component was negatively correlated with the hospital component – evidence of negative assortative matching.

3.5 Describing the Adopters

In this subsection I first present the ex ante relationships one might expect between hospital characteristics and productivity based on theory and prior literature. I then show how these correlations are borne out in my data, with the caveat that these results are descriptive, not causal.

3.5.1 Potential Roles of Hospital Characteristics and Productivity

Size (Number of Beds) Larger hospitals may be more likely to adopt detailed HF coding if there are fixed costs of adoption – fixed costs are smaller, and more worthwhile to incur, on a per-patient basis when the hospital is larger. However, since size may be confounded with other factors that bear on coding, this explanation is only suggestive. In particular, a long line of research has documented a strong relationship between hospital size and quality in many areas, though with an unclear causal link (this is usually called the volume-outcomes hypothesis; see Epstein, 2002 for a critical review).

Ownership The relationship between hospital ownership and coding straddles two broad strands of literature: one that investigates differences in the quality of care by ownership, and another that looks at ownership and the responsiveness to billing incentives. With respect to quality of care, there is no consensus on whether non-profit or for-profit hospitals are superior (McClellan and Staiger, 2000; Sloan, 2000), though for-profit hospitals have lagged public and non-profit facilities
in the use of standards of care like β-blockers (Sloan et al., 2003). The disparities have been clearer in studies of billing and coding, which have found that for-profit hospitals exploited revenue-making opportunities more aggressively than their non-profit and government-run counterparts (Dafny, 2005; Silverman and Skinner, 2004). A key difference between this setting and the earlier work is that the prior literature focused on upcoding, or the exaggeration of patient severity to raise payments. In contrast, achieving a high HF coding rate does not require a hospital to risk the fraud allegations that upcoding can bring. In theory, a hospital can provide a detailed HF code for all its HF patients with detailed documentation but no upcoding.

**Location**  Whether rural hospitals should be more effective at adopting the revenue-raising technology than urban hospitals is unclear *ex ante*, though evidence on outcomes and processes along the dimension of hospital location may be suggestive. There is substantial research indicating that health care outcomes and quality of care are lower in rural hospitals relative to their urban counterparts. At least some of this difference can be explained by rural hospitals being smaller. Hospitals in the farthest outlying rural areas appear to be the main driver of rural hospital underperformance (MEDPAC, 2012b; Baldwin et al., 2010).

**Teaching Status**  Teaching hospitals have been found to have better outcomes and higher quality processes of care than non-teaching hospitals in observational studies (see Ayanian and Weissman, 2002 for a review). These studies do not necessarily control for hospital size, ownership, and other attributes. Still, teaching hospitals appear to be regarded in conventional wisdom as purveyors of the frontier of high quality care (see, for example, *U.S. News and World Report* rankings of hospitals). Whether this conventional wisdom is true, and whether it translates into more responsiveness to incentives in the form of takeup of the revenue-generating practice, is an open question.

**Revenue at Stake**  A hospital with more revenue at stake from the reform, all else equal, would have a greater incentive to buy software that improves specific coding and to coax its doctors to provide detailed documentation. However, the revenue at stake depends on the hospital’s patient mix – hospitals with more HF patients and hospitals with more acute HF patients have more to gain. Even after controlling for a host of observables about the hospitals, unobserved characteristics may still exert an effect on adoption along this gradient.
**Treatment Productivity and Quality** Whether high treatment productivity hospitals are more likely to adopt the coding technology is not obvious. High productivity hospitals may have high quality managers who effectively work with physicians to incorporate consensus standards of care. These managers may use the same techniques to extract more detailed descriptions from their physicians. The managers could also use their treatment productivity-raising techniques to ensure that coding staff does not miss revenue-making opportunities.

On the other hand, a negative correlation between treatment and revenue productivity is also plausible. To the extent that productivity depends on managerial quality, the relationship between revenue productivity and treatment productivity could reflect whether one is a substitute for another in the hospital management production process. In the substitutes view, managers specialize in either coaxing physicians and staff to extract revenue from payers or in pushing them to treat patients well.

### 3.5.2 Results

Table 6 displays the key estimates of the role of hospital characteristics and productivity in explaining takeup of the coding technology. The columns of this table show the results when different sets of first-step controls are included, repeating the sets of controls used in the dispersion analysis.

**Without Physician Controls** Columns (1) to (3) depict the correlations with increasingly rich patient controls, but no physician controls. Column (1), which includes no patient-level adjustments, shows how the raw probability a HF patient at the hospital is billed with a detailed code depends on hospital characteristics. However, these relationships could depend on some hospitals having patients that are harder or easier (or more worthwhile) to code. To address this concern, the next two columns add patient-level risk adjusters. Column (2) shows how hospital characteristics are correlated with the probability that the hospital uses a specific code for a HF patient, first adjusting these probabilities to remove the effect of age, race, sex, source of admission, and main diagnosis. Column (3) adds adjustments for patients’ chronic conditions.

There is a robust relationship between coding and hospital size, non-profit status, use of standards of care, and heart attack treatment productivity. Hospitals that are 10% larger give 0.21pp more of their HF patients a specific code. Adding patient controls when estimating hospital adop-
tion reduces this effect to 0.15pp – some of the raw relationship between size and coding can be accounted by larger hospitals tending to have patients that are more likely to receive a detailed code at any hospital. Likewise, non-profit hospitals give 3.3pp more of their patients a specific code than for-profits and government-run facilities, though adding patient controls reduces the difference to 2.7pp. There is no significant difference between the takeup rates of for-profit and government-run hospitals. Finally, with the full set of patient controls, for each standard deviation rise in heart attack treatment productivity or in the use of standards of care (the composite measure is the sum of the heart attack, heart failure, pneumonia, and surgeries measures), about 2.4pp more HF patients tend to get a specific code. In other words, hospitals that appear to be higher quality and more productive in their treatment are also more likely to use these high-revenue billing codes. Hospital location, teaching status, and revenue at stake are not robustly correlated with takeup.

**With Physician Controls** Columns (4) to (6) repeat the results of columns (1) to (3) with first-step physician controls, changing the interpretation of the coefficients. In these columns, a positive relationship between a hospital characteristic and coding indicates that the facility was able to extract more detailed coding out of its physicians – the hospital effect on the left-hand side of these regressions conditions on the physicians that treated the patients. In this section I focus on the coefficients of column (6), which adjust for the full set of patient characteristics as well as the physician when estimating the hospital component of adoption.

The inclusion of physician effects makes the hospital effects noisier, adding left-hand side measurement error to the regressions. This measurement error comes from sampling variance, so it does not bias the coefficients reported in columns (4) to (6), but it lowers the precision of the regression coefficients.

The use of detailed HF codes is clearly correlated with both heart attack treatment productivity and the use of consensus standards of care: even if all hospitals had the same kinds of patients and doctors, hospitals with one standard deviation greater use of standards of care or one standard deviation greater treatment productivity would use specific codes for 2-3pp more of their patients. This gradient was also observed unconditional on the doctors (in columns 1-3), and these results indicate that it cannot be explained by high treatment quality hospitals simply having physicians that provide detailed documentation wherever they practice. Instead, these results indicate that
these hospitals are more able to extract the codes from their physicians than their lower treatment quality peers.

The other significant relationship that exists with first-step patient and physician controls is that between hospital location and coding. Hospitals in large urban and other urban areas – areas of high and intermediate population density, respectively – extract specific codes from their doctors for 3-4pp more of their patients than hospitals in rural areas. This relationship does not exist without the physician controls, which indicates that urban hospitals have physicians that are less likely to provide detailed documentation wherever they work, but that the low physician contribution is counteracted by the hospitals’ ability to extract codes from their doctors. The net result is that unconditional on physicians, urban and rural hospitals are about equally likely to use the detailed billing codes – the finding in columns (1)-(3).

Non-profit and for-profit hospitals were 1.7pp more likely to extract specific codes from doctors than their government-run counterparts, though these coefficients were imprecisely measured. Compared to the same differential calculated unconditional on physicians – the result in column (3) – this value is reduced and no longer significant for the non-profit hospitals. Since removing the physician component of adoption reduces the coding advantage of non-profit facilities, it appears that the physicians who work at non-profit hospitals are more likely to provide the detailed documentation wherever they practice.

The gradient between hospital size and extraction of detailed HF codes is positive and significant without first-step physician controls, but it is eliminated when the physician component of adoption is swept away. Similar to the results for non-profit hospitals, this finding suggests that larger hospitals outperform smaller hospitals in column (3) because they utilize physicians that always provide more documentation wherever they treat patients.

4 Discussion

The hallmark features of a new technology are wide variations in the level of adoption at a point in time and variation in adoption over time as takeup slowly occurs. This pattern is found in Griliches (1957), and it has also been found in health care, for example in the use of β-blockers and other therapies (see e.g. Bradley et al., 2005 and Peterson et al., 2008). Likewise, a growing literature
is finding persistent dispersion in productivity within narrowly defined industries (Fox and Smeets, 2011; Syverson, 2011); this literature is now expanding to include the health care sector (Chandra et al., 2013). I have shown that adoption of the HF coding technology across hospitals follows the established pattern.

Some hospitals may be very detailed coders because their doctors are likely to provide specific documentation wherever they practice. Other hospitals might take up the revenue generating practice by counteracting the poor documentation habits of their physicians with facility-specific techniques, like aggressively reviewing physician charts. Uniquely in the HF coding setting I can observe the component of adoption that is specific to the hospital – the extent to which a hospital can extract more details out of a constant set of physicians than other hospitals.

Since hospitals but not physicians were paid for the HF documentation, I have argued that the hospital component of adoption is an indicator for whether the hospital was able to solve a principal-agent problem. This component is robustly correlated with the use of consensus standards of care when treating patients. Thus hospitals that use treatment productivity-raising techniques are able to extract more specific documentation from a fixed set of physicians than other hospitals. The correlation between these two measures suggests that agency problems could play a role in the adoption of a variety of technologies in the facility. Another view of this correlation is that revenue productivity and treatment productivity are positively related.

The dispersion that I find in the hospital component of adoption, which removes the physician and patient components, is about four-fifths the raw level of dispersion. This residual dispersion has a standard deviation of 0.16 percentage points, but it is not immediately clear whether this magnitude is small or large. One point of comparison is the standard deviation of the consensus standards of care scores, which measure adherence to evidence-based treatment guidelines. The measures of coding of HF and standards of care are both hospital-level shares, so it is reasonable to compare their variances. To the extent that there are substantial disparities across hospitals in their adherence to these standards, the disparities in coding also appear to be nontrivial. According to Table 4, the four standards of care scores have standard deviations ranging from 0.07pp to 0.14pp. The dispersion in the hospital component of HF coding adoption falls just above the top end of this range.

As public insurers move to incentivize the adoption of consensus health care treatments, the
effects that these incentives will have remain unclear. Looking at the relationships between HF
coding and hospital characteristics sheds light both on the likely effects of future incentives as well
as the mechanisms that drive incomplete takeup. In particular, these correlates offer evidence on
which providers are likely to be policy elastic to financial incentives for other processes of care. For
the policy elasticity, it is useful to look at the correlation between takeup and characteristics without
removing the effect of the physician, since the overall response of the hospital is of interest. I have
shown that bigger, non-profit, higher treatment quality, and more treatment-productive hospitals
are more policy elastic.

One reason to incentivize the use of evidence-based inexpensive medical technologies is to push
lagging hospitals to take them up. Quality disparities have been a key focus of health care literature
(see e.g. Fisher et al., 2003), and policymakers are increasingly using direct financial incentives with
the hope of improving outcomes at low-performing hospitals. For example, the Medicare Value-
Based Purchasing program will reduce payments to hospitals that fail to use consensus standards
of care or whose patients report low satisfaction with their experiences. Yet it is an open question
whether these policies will have their intended effect of raising quality; according to these findings,
policy elastic providers tend to be more productive in treatment and more likely to follow consensus
standards of care already. Lower quality providers – i.e. those that are less productive or less
likely to follow best practices – are less responsive. These results suggest that hospitals that are
behind the curve on medical standards are also less attuned to financial incentives, which means
that policies to incentivize takeup could have their least effect on the providers that need the most
improvement. In turn, these programs could serve to widen disparities in the quality of care across
providers.

5 Conclusion

This paper has examined the takeup of a revenue-generating practice – the use of specific, detailed
codes to describe heart failure on inpatient claims – that was incentivized following a 2008 reform.
I have shown that hospitals responded by rapidly improving the documentation of patients in their
claims. Yet this improvement in documentation was incomplete and uneven, a characteristic feature
of the adoption of new technologies. I have also decomposed the takeup of the technology into a
component that is due to the hospital and a component that is due to its doctors. The decomposition exercise shows that hospitals that had high treatment productivity and followed consensus standards of care were better able to extract detailed documentation from their physicians. I argue that this is consistent with these hospitals solving principal-agent problems.

My results have important policy implications as public and private insurers seek to directly raise hospital productivity by reforming health care payment systems. Principal-agent problems owing to a bifurcated system that pays doctors and hospitals on separate bases may be major impediments to further technology adoption. For example, when Medicare opts to pay hospitals to use $\beta$-blockers, it trusts that the facilities will recognize the financial gains to changing their processes of care and successfully transmit the incentives to the physicians who prescribe the drugs. Yet some facilities appear much more able to transmit these incentives than others.

One potential policy to obviate the incentive transmission problem is to reform the physician payment system. Provisions of the Affordable Care Act that require this system to incentivize standards of care, much as Medicare is already doing for hospital payments, are one way forward. By bringing these incentives to both hospitals and doctors, these provisions could substantially improve the effectiveness of value-based payment reforms.

References


Baldwin, Laura-Mae, Leighton Chan, C .Holly A. Andrilla, Edwin D. Huff, and L. Gary Hart. 2010. “Quality of Care for Myocardial Infarction in Rural and Urban HospitalsQuality of


Figures

Diffusion of Coding Technology Over Time

Figure plots the share of revenue available for detailed coding of HF that was captured by hospitals over time. Dotted line shows revenue that would have been captured in 2007 if hospitals had been paid per 2008 rules. The series is at the weekly level and the red line denotes the reform date.

Figure 1

Organizational Process for Coding

Doctor writes down patient’s diagnoses → Patient’s Medical Chart

Hospital’s coding staff queries doctor → Updated Chart

Coding staff translates chart to diagnosis codes → Claim to be sent to Medicare

Figure 2
Gain in Revenue by Type of Detailed HF Code

![Gain in Revenue by Type of Detailed HF Code](image1)

Figure plots the average per-HF patient gain in revenue going from always using vague codes for HF patients to always using chronic codes or acute codes.

**Figure 3**

Use of Detailed HF Codes Over Time

![Use of Detailed HF Codes Over Time](image2)

Figure plots the share of HF patients who received a detailed HF code over time. The series is at the weekly level and the red line denotes the reform date.

**Figure 4**
Revenue at Stake per HF Patient across Hospitals

Hospitals: 3,103
Mean: $1,006.90
SD: $229.60

Revenue at stake is calculated using pre-reform (2007) patients processed under post-reform (2009) payment rules. The amount at stake equals the per-HF patient revenue with all HF patients given detailed codes less that revenue with all HF patients given vague codes. The 422 hospitals with <50 HF patients are suppressed. The outlying upper and lower 1% of hospitals were also suppressed.

Figure 5

Revenue at Stake per Patient across Hospitals

Hospitals: 3,103
Mean: $267.64
SD: $76.14

Revenue at stake is calculated using pre-reform (2007) patients processed under post-reform (2009) payment rules. The amount at stake equals the per-patient revenue with all HF patients given detailed codes less that revenue with all HF patients given vague codes. The 422 hospitals with <50 HF patients are suppressed. The outlying upper and lower 1% of hospitals were also suppressed.

Figure 6
**Figure 7**

HF Coding and Heart Testing Following Reform

![Graph showing detailed coding and cardiac echo testing over years](image)

- **Detailed Coding** (Green line)
- **Cardiac Echo Testing** (Red dashed line)

Figure plots the weekly share of revenue available for detailed coding of HF that was captured by hospitals alongside the weekly share of all patients who received a cardiac echo, a heart test. The dotted line shows revenue that would have been captured in 2007 if hospitals had been paid per 2008 rules. The red line denotes the reform date.

**Figure 8**

Distribution of Adoption across Hospitals over Time

![Box and whisker plots showing adoption distribution](image)

- **Box and whiskers** show the distribution of adoption across hospitals in each year. A hospital’s adoption equals the share of its HF patients in that year who received a specific HF code.
- Hospital-years with fewer than 50 HF patients are excluded. Red line separates pre- and post-reform years.

Box and whiskers show the distribution of adoption across hospitals in each year. A hospital’s adoption equals the share of its HF patients in that year who received a specific HF code. Hospital-years with fewer than 50 HF patients are excluded. Red line separates pre- and post-reform years.
A hospital’s adoption equals the share of its 2010 HF patients who received a detailed HF code. Hospitals with fewer than 50 HF patients in 2010 (N=441) are excluded.

Figure 9
### Table 1 - Vague and Specific HF Codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Severity Before</th>
<th>Severity After</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Vague Codes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>428.0</td>
<td>Congestive HF, Unspecified</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>428.9</td>
<td>HF, Other</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td><strong>Specific Codes (Exhaustive Over Types of HF)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>428.20</td>
<td>HF, Systolic, Onset Unspecified</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>428.21</td>
<td>HF, Systolic, Acute</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>428.22</td>
<td>HF, Systolic, Chronic</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>428.23</td>
<td>HF, Systolic, Acute on Chronic</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>428.30</td>
<td>HF, Diastolic, Onset Unspecified</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>428.31</td>
<td>HF, Diastolic, Acute</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>428.32</td>
<td>HF, Diastolic, Chronic</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>428.33</td>
<td>HF, Diastolic, Acute on Chronic</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>428.40</td>
<td>HF, Combined, Onset Unspecified</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>428.41</td>
<td>HF, Combined, Acute</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>428.42</td>
<td>HF, Combined, Chronic</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>428.43</td>
<td>HF, Combined, Acute on Chronic</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Congestive HF (the description of code 428.0) is often used synonymously with HF. Other HF codes include 428.1 (Left HF); 398.91 (Rheumatic HF); and 402.x1, 404.x1, and 404.x3 (forms of hypertension alongside HF). These other codes were all high-severity before the reform and most of them remained medium- or high-severity after the reform. The only exceptions were two codes that could be used alongside a specific code to maintain a medium or high level of severity following the reform.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>HF Patients</td>
<td>552.7</td>
<td>578.5</td>
<td>1</td>
<td>5,435</td>
</tr>
<tr>
<td>HF Patients (mobility sample)</td>
<td>581.7</td>
<td>553.1</td>
<td>1</td>
<td>4,607</td>
</tr>
<tr>
<td>Distinct Physicians</td>
<td>57.9</td>
<td>54.1</td>
<td>1</td>
<td>561</td>
</tr>
<tr>
<td>Mobile Physicians</td>
<td>19.8</td>
<td>21.9</td>
<td>1</td>
<td>173</td>
</tr>
</tbody>
</table>

Hospitals (N=3,414; N=2,868 in mobility sample)

Physicians (N=134,502)

Italicsized rows refer to the mobility sample: the subset of the analysis sample in which I observe the physician and can separately identify the hospital and physician effects. See text for more details. Full sample includes 1.9M HF patients. Mobility sample includes 1.7M HF patients.
Table 3 - Statistics about Physicians by Mobility Status

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All values are means</td>
<td>All</td>
<td>Mobile</td>
<td>Non-Mobile</td>
</tr>
<tr>
<td>Patient and Hospital Volume</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF Patients</td>
<td>12.4</td>
<td>21.6</td>
<td>10.3</td>
</tr>
<tr>
<td>Distinct Hospitals</td>
<td>1.23</td>
<td>2.25</td>
<td>1</td>
</tr>
<tr>
<td>Mobile (&gt;1 hospital)</td>
<td>0.19</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Type of Physician</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary Physician</td>
<td>0.44</td>
<td>0.50</td>
<td>0.43</td>
</tr>
<tr>
<td>Medical Specialist</td>
<td>0.30</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Surgeon</td>
<td>0.23</td>
<td>0.14</td>
<td>0.25</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.19</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>Age</td>
<td>49.1</td>
<td>48.9</td>
<td>49.1</td>
</tr>
<tr>
<td>Training and Experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years in Training</td>
<td>5.96</td>
<td>6.52</td>
<td>5.83</td>
</tr>
<tr>
<td>Years Since Training</td>
<td>16.0</td>
<td>15.4</td>
<td>16.1</td>
</tr>
<tr>
<td>Trained in US</td>
<td>0.71</td>
<td>0.59</td>
<td>0.74</td>
</tr>
<tr>
<td>Physicians</td>
<td>134,502</td>
<td>25,253</td>
<td>109,249</td>
</tr>
</tbody>
</table>

Mobile physicians are observed attending to HF patients at multiple hospitals in 2010; non-mobile physicians attend to patients at one hospital in that period. Data on physician type, demographics, training, and experience derived from AMA Masterfile.
### Table 4 - Hospital Summary Statistics

<table>
<thead>
<tr>
<th>Patient Controls</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2,411</td>
<td>709.4</td>
<td>606.3</td>
</tr>
<tr>
<td>Mean</td>
<td>2,411</td>
<td>0.546</td>
<td>0.201</td>
</tr>
<tr>
<td>SD</td>
<td>2,411</td>
<td>285.0</td>
<td>231.4</td>
</tr>
<tr>
<td>Ownership</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Profit</td>
<td>2,411</td>
<td>0.667</td>
<td>0.471</td>
</tr>
<tr>
<td>For-Profit</td>
<td>2,411</td>
<td>0.164</td>
<td>0.371</td>
</tr>
<tr>
<td>Government</td>
<td>2,411</td>
<td>0.169</td>
<td>0.375</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Urban Area</td>
<td>2,411</td>
<td>0.419</td>
<td>0.494</td>
</tr>
<tr>
<td>Other Urban Area</td>
<td>2,411</td>
<td>0.350</td>
<td>0.477</td>
</tr>
<tr>
<td>Rural Area</td>
<td>2,411</td>
<td>0.231</td>
<td>0.421</td>
</tr>
<tr>
<td>Teaching Hospital</td>
<td>2,411</td>
<td>0.371</td>
<td>0.483</td>
</tr>
<tr>
<td>Ex Ante $ at Stake / Patient</td>
<td>2,411</td>
<td>268.6</td>
<td>72.22</td>
</tr>
<tr>
<td>Standards of Care (share of times standards were used in 2006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for Heart Attack Treatment</td>
<td>2,411</td>
<td>0.916</td>
<td>0.085</td>
</tr>
<tr>
<td>for Heart Failure Treatment</td>
<td>2,411</td>
<td>0.826</td>
<td>0.113</td>
</tr>
<tr>
<td>for Pneumonia Treatment</td>
<td>2,411</td>
<td>0.864</td>
<td>0.061</td>
</tr>
<tr>
<td>for High-Risk Surgeries</td>
<td>2,411</td>
<td>0.797</td>
<td>0.119</td>
</tr>
<tr>
<td>ln(Productivity)</td>
<td>2,411</td>
<td>0.919</td>
<td>0.171</td>
</tr>
</tbody>
</table>

A large urban area is an MSA with a population of at least 1 million; the remaining MSAs are considered other urban areas. A rural area is any location outside an MSA. A hospital's ex ante $ at stake per patient is the revenue put at stake by the reform per patient in the hospital (including non-HF patients). See text for more details on the standards of care and heart attack treatment measures. The standard deviation of heart attack treatment productivity is adjusted for sampling variance.
Table 5 - Standard Deviation of Coding by Type of Hospital

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Controls</strong></td>
<td>None</td>
<td>Admission</td>
<td>Full</td>
<td>None</td>
<td>Admission</td>
<td>Full</td>
</tr>
<tr>
<td><strong>Physician Controls</strong></td>
<td>None</td>
<td>Physician Fixed Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Urban Non-Profit Hospitals</strong></td>
<td>0.168</td>
<td>0.135</td>
<td>0.134</td>
<td>0.145</td>
<td>0.116</td>
<td>0.123</td>
</tr>
<tr>
<td>Large and Teaching</td>
<td>[281]</td>
<td>[281]</td>
<td>[281]</td>
<td>[281]</td>
<td>[281]</td>
<td>[281]</td>
</tr>
<tr>
<td>Other</td>
<td>0.187</td>
<td>0.142</td>
<td>0.141</td>
<td>0.193</td>
<td>0.158</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>[407]</td>
<td>[407]</td>
<td>[405]</td>
<td>[407]</td>
<td>[407]</td>
<td>[405]</td>
</tr>
<tr>
<td><strong>Non-Urban Non-Profit Hospitals</strong></td>
<td>0.167</td>
<td>0.134</td>
<td>0.134</td>
<td>0.223</td>
<td>0.191</td>
<td>0.193</td>
</tr>
<tr>
<td>Teaching</td>
<td>[316]</td>
<td>[316]</td>
<td>[316]</td>
<td>[316]</td>
<td>[316]</td>
<td>[316]</td>
</tr>
<tr>
<td>Other</td>
<td>0.211</td>
<td>0.161</td>
<td>0.159</td>
<td>0.200</td>
<td>0.167</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>[604]</td>
<td>[604]</td>
<td>[593]</td>
<td>[604]</td>
<td>[604]</td>
<td>[593]</td>
</tr>
<tr>
<td><strong>For-Profit and Government-Run Hospitals</strong></td>
<td>0.194</td>
<td>0.146</td>
<td>0.145</td>
<td>0.188</td>
<td>0.144</td>
<td>0.135</td>
</tr>
<tr>
<td>Urban For-Profit</td>
<td>[167]</td>
<td>[167]</td>
<td>[167]</td>
<td>[167]</td>
<td>[167]</td>
<td>[167]</td>
</tr>
<tr>
<td>Non-Urban For-Profit</td>
<td>0.188</td>
<td>0.141</td>
<td>0.140</td>
<td>0.178</td>
<td>0.158</td>
<td>0.156</td>
</tr>
<tr>
<td>Government-Run</td>
<td>[229]</td>
<td>[229]</td>
<td>[224]</td>
<td>[229]</td>
<td>[229]</td>
<td>[224]</td>
</tr>
<tr>
<td>All Hospitals</td>
<td>0.222</td>
<td>0.163</td>
<td>0.161</td>
<td>0.177</td>
<td>0.146</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>[407]</td>
<td>[407]</td>
<td>[400]</td>
<td>[407]</td>
<td>[407]</td>
<td>[400]</td>
</tr>
<tr>
<td></td>
<td>[2,411]</td>
<td>[2,411]</td>
<td>[2,386]</td>
<td>[2,411]</td>
<td>[2,411]</td>
<td>[2,386]</td>
</tr>
</tbody>
</table>

Numbers in brackets count hospitals used to calculate the standard deviation. All results are adjusted for sampling variation. Urban hospitals are defined as those located in a "Large Urban" area. Large hospitals are defined as having at least 250 beds. Columns 1 and 4 include no patient controls when calculating the hospital's coding score. Columns 2 and 5 control for patient age, race, sex, admission through the emergency department, and principal diagnosis category. Columns 3 and 6 add controls for histories of chronic conditions. Columns 4-6 control for physician fixed effects when calculating the hospital's coding score.
Table 6 - Describing the Distribution of Coding with Hospital Characteristics and Productivity

<table>
<thead>
<tr>
<th>Patient Controls</th>
<th>None</th>
<th>Admission</th>
<th>Full</th>
<th>None</th>
<th>Admission</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Beds)</td>
<td>0.0214***</td>
<td>0.0146**</td>
<td>0.0145**</td>
<td>0.00603</td>
<td>-0.000296</td>
<td>-0.00256</td>
</tr>
<tr>
<td></td>
<td>(0.00775)</td>
<td>(0.00604)</td>
<td>(0.00599)</td>
<td>(0.0117)</td>
<td>(0.0100)</td>
<td>(0.0104)</td>
</tr>
<tr>
<td>ln(Beds)</td>
<td>0.0328**</td>
<td>0.0250**</td>
<td>0.0269**</td>
<td>0.0194</td>
<td>0.0173</td>
<td>0.0172</td>
</tr>
<tr>
<td></td>
<td>(0.00604)</td>
<td>(0.00599)</td>
<td>(0.0117)</td>
<td>(0.0100)</td>
<td>(0.0104)</td>
<td></td>
</tr>
<tr>
<td>ln(Beds)</td>
<td>0.0168</td>
<td>0.0124</td>
<td>0.0124</td>
<td>0.0198</td>
<td>0.0165</td>
<td>0.0172</td>
</tr>
<tr>
<td></td>
<td>(0.0118)</td>
<td>(0.0117)</td>
<td>(0.0117)</td>
<td>(0.0187)</td>
<td>(0.0196)</td>
<td></td>
</tr>
<tr>
<td>ln(Beds)</td>
<td>0.0830</td>
<td>0.00352</td>
<td>0.00178</td>
<td>0.0627***</td>
<td>0.0446**</td>
<td>0.0370*</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.00809)</td>
<td>(0.00804)</td>
<td>(0.0123)</td>
<td>(0.0130)</td>
<td></td>
</tr>
<tr>
<td>ln(Beds)</td>
<td>0.0110</td>
<td>0.0125</td>
<td>0.0103</td>
<td>0.0632***</td>
<td>0.0406**</td>
<td>0.0332**</td>
</tr>
<tr>
<td></td>
<td>(0.0140)</td>
<td>(0.0105)</td>
<td>(0.0106)</td>
<td>(0.0190)</td>
<td>(0.0167)</td>
<td>(0.0168)</td>
</tr>
<tr>
<td>Teaching Hospital</td>
<td>0.00854</td>
<td>0.0118</td>
<td>0.0112</td>
<td>-0.00531</td>
<td>0.0111</td>
<td>0.0144</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.00809)</td>
<td>(0.00804)</td>
<td>(0.0144)</td>
<td>(0.0123)</td>
<td>(0.0130)</td>
</tr>
<tr>
<td>Teaching Hospital</td>
<td>6.67e-05</td>
<td>6.48e-05</td>
<td>7.64e-05</td>
<td>9.34e-05</td>
<td>8.85e-05</td>
<td>0.000108</td>
</tr>
<tr>
<td></td>
<td>(6.41e-05)</td>
<td>(4.82e-05)</td>
<td>(4.66e-05)</td>
<td>(8.90e-05)</td>
<td>(8.17e-05)</td>
<td>(8.64e-05)</td>
</tr>
<tr>
<td>Ex Ante $ at Stake per Patient</td>
<td>0.0336***</td>
<td>0.0244***</td>
<td>0.0247***</td>
<td>0.0247***</td>
<td>0.0209***</td>
<td>0.0225***</td>
</tr>
<tr>
<td></td>
<td>(0.00559)</td>
<td>(0.00428)</td>
<td>(0.00432)</td>
<td>(0.00691)</td>
<td>(0.00624)</td>
<td>(0.00632)</td>
</tr>
<tr>
<td>Heart Attack Treat Productivity Z-Score</td>
<td>0.0288***</td>
<td>0.0240***</td>
<td>0.0237***</td>
<td>0.0309***</td>
<td>0.0305***</td>
<td>0.0272***</td>
</tr>
<tr>
<td></td>
<td>(0.00629)</td>
<td>(0.00479)</td>
<td>(0.00485)</td>
<td>(0.00863)</td>
<td>(0.00760)</td>
<td>(0.00776)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,411</td>
<td>2,411</td>
<td>2,386</td>
<td>2,411</td>
<td>2,411</td>
<td>2,386</td>
</tr>
<tr>
<td>$R^2$ (adjusted)</td>
<td>0.091</td>
<td>0.091</td>
<td>0.092</td>
<td>0.037</td>
<td>0.036</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Standard errors clustered at the market level in parentheses. Columns 1 and 4 include no patient controls when calculating the hospital's coding score. Columns 2 and 5 control for patient age, race, sex, admission through the emergency department, and principal diagnosis category. Columns 3 and 6 add controls for histories of chronic conditions. Columns 4-6 control for physician fixed effects when calculating the hospital's coding score. The standards of care composite z-score is the sum of the four standards of care measures, normalized to mean 0 and standard deviation 1.

*** significant at 1% level; ** significant at 5% level; * significant at 10% level