Shall We Blame Health IT for Medicare Overpayments? New Evidence from Medicare Recovery Audit Program

Keith A. Joiner * Jianjing Lin †

Abstract

We study how the adoption of electronic medical records (EMRs) affects Medicare payments for hospitalization. Increasing Medicare bills for inpatient care associated with the rapid diffusion of EMRs can arise from the following two mechanisms: (1) EMRs inflate Medicare bills to a higher level than justified, and (2) EMRs facilitate complete documentation to recover more billable services than before. Prior studies based on claims data may not distinguish between these two mechanisms, as both of them lead to similar patterns. Using data on audit outcomes from a national audit program, we examine how the amount of overpayments / underpayments for hospitalization changes as a hospital adopts EMRs. We find no significant correlation between EMR adoption and overpayments, while underpayments decrease with EMR adoption, supporting the complete-coding mechanism.

JEL Codes: H51, I13, I18, O33
Keywords: Healthcare expenditure, improper billing, audits, health IT, hospital inpatient care

*University of Arizona

†Rensselaer Polytechnic Institute; Email: linj17@rpi.edu

1 Introduction

The rapid diffusion of electronic medical records (EMRs)—the central component of health information technology (IT)—in the last decade is partly due to the aggressive promotion from the U.S. federal government as well as its potential in improving productivity and quality of care. However, the fundamental reason for healthcare providers to adopt this technology is to optimize reimbursements, especially given the complexity in medical billing. The widespread adoption of EMRs is described as a contributor to the growth of Medicare bills. This could occur in the following two fashions: (1) the *bill-inflation* mechanism, where EMRs facilitate healthcare providers submitting bills that are not medically justifiable in response to financial incentives ("upcoding");[1] and (2) the *complete-coding* mechanism, where EMRs enable providers to code more completely so as to recover more billable services than before (a practice that is called "charge capture"). Prior studies examining claims data may not separate these two mechanisms, as both of them could lead to similar patterns. We revisit this topic using information on audit outcomes from the Medicare Recovery Audit Program, which could help identify which is the dominant mechanism.

To illustrate why this is the situation, we present an example reflecting a common clinical situation leading to hospitalization—heart failure. The reimbursement for inpatient care is based on the recorded diagnostic related group (DRG) and then adjusted to the corresponding severity level—with higher payments for a more severe level—according to the accompanying complicating conditions. Consider a patient admitted to a hospital because of heart failure. In general, one

[1] The Centers for Medicare and Medicaid Services (CMS) provide a definition of upcoding: "Upcoding refers to billing a higher level service or a service with a higher payment than is supported by the medical record documentation" (Supplementary Appendices, CERT Program, 2012).

of the three DRGs that describes heart failure and shock can be assigned to the patient as long as the base diagnosis is selected from the 20 associated ICD-10 codes, that discriminate between functional abnormalities, anatomic abnormalities, and acuity, in various combinations.[2] Differentiation between these codes requires documentation in the EMR, using history (past and present), physical examination, laboratory and imaging tests, response to therapy and more.

Accurate documentation of the disease code of heart failure, as described above, is *only* one component of the final process of coding. Of equal if not more importance in coding and reimbursement is the presence or absence of a complication or comorbidity (CC) or major complication or comorbidity (MCC), modifying the base diagnosis by reflecting greater resource use for management. For example, an individual with acute systolic heart failure might develop a pleural effusion. Pleural effusions occur when fluid collects between the outside lining of the lung and the pleura, which lines the chest cavity. This constitutes a CC. Pleural effusions are typically identified from chest x-rays or other imaging studies of the lungs, such as CT scans, and would be noted on the reports from those procedures. Because pleural effusions are common in this setting, they may not be considered as a separate diagnosis, and consequently be omitted from the diagnosis list. This would be an example of downcoding.[3] Upcoding would occur if pleural effusion was included on the diagnosis list, when it was not present. This might occur if a pleural effusion present on a past hospitalization had resolved, but the diagnosis was not deleted.

---

[2]See `https://www.cms.gov/icd10manual/version33-fullcode-cms/fullcode_cms/P0139.html`. ICD-10 codes stand for International Classification of Diseases codes, 10^{th} version.

[3]Coding refers to a process of translating medical information—that includes diagnosis, procedures, medical services and equipment—into universal medical alphanumeric codes for documentation and medical billing purposes.

Of key importance, analysis of a Medicare claim for the above hospitalization cannot distinguish whether the documentation in the medical record was sufficient or insufficient to justify the diagnosis of a pleural effusion. Neither the clinical nor imaging data suggesting the presence (or absence) of a pleural effusion, nor other features of the clinical course can be verified from the claim. However, the presence of pleural effusion noted in the report from the imaging study will be automatically included in the EMR, not requiring further documentation on the part of the provider or coder. The requested reimbursement could be accurate, or could reflect overpayment or underpayment. The only way to evaluate the accuracy of the information is a direct review of the medical record. For the individual patient, this is done by the coders who prepare the claim for submission.[4] Regardless of the analytical methods used to evaluate the accuracy of coding using large data sets, chart audits are required for direct verification (Bauder et al., 2017).

A common intuition that the primary incentive in coding is to maximize revenue, leading to upcoding, belies the complexity interplay of factors involved.[5] When information is left out or erroneously recorded, providers or coders tend to choose a lower billing code to avoid the scrutiny by insurers or auditors.[6] Sacarny (2018) also finds that hospitals face large frictions in using the appropriate level of billing codes but instead using a more general billing code that leads to lower reimbursements.

Prior studies that examine the correlation between EMR adoption and billing mainly use claims

[4]Healthcare coders are the staff mainly responsible for coding.

[5]See https://www.nytimes.com/2017/03/29/magazine/those-indecipherable-medical-bills-theyre-one-reason-health-care-costs-so-much.html.

[6]See https://www.todayshospitalist.com/Breaking-the-downcoding-habit/.

data (Adler-Milstein and Jha, 2014; Li, 2014; Ganju et al., 2021; Gowrisankaran et al., 2022). However, both the complete-coding and bill-inflation mechanisms result in very similar patterns in claims data—in particular, more reported CCs or MCCs among EMR hospitals.

In this paper, we use information on review decisions of medical claims from the Medicare Recovery Audit Program (RAP), to examine the extent to which EMRs affect improper billing and to distinguish the bill-inflation mechanism from the complete-coding mechanism. RAP is one of the auditing strategies by the Centers for Medicare and Medicaid Services (CMS) to evaluate the accuracy of Medicare payments; almost all Medicare Parts A and B claims are subject to recovery auditors' review.[7] Recovery auditors (RAs) are paid a fraction of identified improper payments and must return the fee if the audit determination is overturned in an appeal. As a result, they have incentives to ensure the accuracy of the review decision.

If the bill-inflation mechanism dominates, we expect more overpayment determinations among EMR hospitals, as RAs are likely to detect inappropriate billing practices, especially after the complex review of all supporting documents.[8] If, instead, the complete-coding mechanism plays a more important role, we expect no significant difference in audit findings (especially for overpayments) between EMR and non-EMR hospitals, as EMRs simply facilitate complete documentation which should comply with the reimbursement criteria.

We are one of the first studies using audit outcomes to understand the underlying effect of

[7] See https://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/Medicare-FFS-Compliance-Programs/Recovery-Audit-Program/Downloads/RAC-SOW-Regions-1-4-clean-November-30-2016.pdf.

[8] Most of the identified improper payments come from the complex review conducted by healthcare coders or clinicians.

EMRs on hospital billing and coding practices. The results of the analysis may lead to different policy implications. The presence of the bill-inflation mechanism may call for greater fraud enforcement to deter improper billing and policy interventions to ensure that electronic documentation is valid, accurate, and complete.[9] Conversely, the finding that the complete-coding mechanism dominates may suggest that compliance with Medicare billing requirements is costly, and EMRs make this process less challenging. Given the difficulty in and importance of complying with the documentation criteria, policy makers may consider offering appropriate guidance and incentives to encourage proper documentation among providers.

In this paper, we study how EMR adoption affects improper payments in hospital inpatient care, including overpayments and underpayments. Based on novel audit-level administrative data from a national audit program—the Recovery Audit Program, we examine how the amount of identified improper payments varies as a hospital adopts EMRs.[10] We apply a fixed effect estimation approach and find little effect of EMR adoption on overpayments and an average decrease of 7.06% in underpayments among EMR hospitals. We use an event study research design to test whether the change in improper payments among EMR hospitals prior to their adoption was systematically different from those among non-EMR hospitals and find that there were no differential pre-adoption trends in improper payments between these two types of hospitals.

To explore the underlying mechanisms, we also estimate the effects by RAs' capabilities. Cer-

[9] Such policy interventions may include rules or guidance on authorship validation, document amendments, auditing the record for documentation validity, and so on.

[10] We thank Maggie Shi for sharing these data, which is used in Shi (2021). She obtained these data via a Freedom of Information Act request.

tain RAs have developed superior capabilities of detecting the use of copy-and-paste entries and inappropriate documentation in EMRs. We expect RAs with such capabilities are more likely to identify overpayments if the bill-inflation mechanism dominates. Our results indicate no effects of EMRs on overpayments, regardless of the RA's capabilities, but the reduction in underpayments as a consequence of EMR adoption concentrates on claims reviewed by RAs with the specialized capabilities. In our heterogeneity analysis, we find that the significant effect on underpayments mainly occurs with not-for-profit hospitals or financially distressed hospitals. The finding of little impact on overpayments along with lower underpayments from EMR adoption is somewhat consistent with the finding in previous studies that EMRs result in greater revenue, but through charge capture rather than upcoding.

Related literature. Our paper is related to three strands of literature. First, it contributes to the research studying the effect of EMR adoption on hospital billing and coding practices. This literature has found mixed results. Adler-Milstein and Jha (2014) find that EMR hospitals do not increase billing to Medicare to a greater extent than non-EMR hospitals, whereas Li (2014) and Ganju et al. (2021) find relatively more patients reported with severe conditions in EMR hospitals. Another paper by Qi and Han (2020) examines the effect of a broader set of health ITs and find that the adoption of these technologies boosts patient revenue. Of importance, none of these papers distinguishes between the bill inflation and complete coding mechanisms. Our paper complements these studies by separating the two mechanisms.

Second, our paper contributes to the broader literature on hospital upcoding behavior.[11] An

[11] There are other studies examining bill inflation/upcoding in settings other than hospitals, such as Brunt (2011), Bowblis and Brunt (2014), Fang and Gong (2017) and Geruso and Layton (2020).

influential paper by Dafny (2005) leverages an exogenous change in DRG price and finds that hospitals coded more patients to the severe category when the reimbursement increment was larger. Silverman and Skinner (2004) find that for-profit hospitals experienced the largest percentage increase in patients assigned to the most generous DRG for pneumonia and respiratory infections, which is aligned with the goal of the administration in these hospitals. A more recent paper by Cook and Averett (2020) find that approximately three percent of reimbursements could be attributed to upcoding after the DRG restructuring in 2008. Such findings are not unique in the U.S. healthcare system (Jürges and Köberlein, 2015; Januleviciute et al., 2016; Barros and Braun, 2017; Verzulli et al., 2017). While it is inarguable that hospitals have incentives for enhancing revenues, the observation of increasing Medicare bills could also arise from a reduction in down-coding, due to improved efficiency in documentation and billing. Our paper contributes to this literature by using the data from a national audit program, in the hope of identifying the underlying mechanism in the more recent context.[12]

Finally, our paper also contributes to a small strand of literature that explores the cost of complete coding. A related study by Sacarny (2018) finds that hospitals sometimes pick a generic code for heart failure instead of specifying the type, even though any specific code could generate greater reimbursements, which may imply the cost of complete documentation. In other words, hospitals may down-code because the cost of complete coding is substantial.

The rest of the paper proceeds as follows. Section 2 describes industry and institutional background. Section 3 introduces the datasets and reports summary statistics. Section 4 presents the

[12]Our ongoing work also aims to separate the two mechanisms but using Medicare claims data and has similar findings to this paper (Gowrisankaran et al., 2022).

empirical strategy. Section 5 discusses the estimation results. The last section concludes.

2 Background

2.1 MEDICARE INPATIENT BILLING AND THE ROLE OF EMRS

Hospital inpatient admissions are paid based on a flat-rate payment system, known as the inpatient prospective payment system (IPPS). A hospital assigns a single DRG for each patient stay based on the primary and additional diagnoses/procedures. Each DRG is associated with a weight that reflects average resources used to treat Medicare patients in that DRG. Medicare reimburses the hospital a flat rate for the admission, which is proportional to the DRG weight plus some adjustments reflecting hospital and region specifics. Typically, the base DRG is first determined, according to the patient's primary diagnosis/procedure. It is likely that several DRGs share a common base DRG. For instance, DRGs 637 – 639 are "Diabetes w MCC," "Diabetes w CC," and "Diabetes w/o CC/MCC," respectively. All three belong to the same base DRG—diabetes. The base DRG can be modified depending on the presence or absence of a CC or MCC, which increases the weight of the DRG and hence the reimbursement.

Justification for selecting a specific DRG, with or without a CC or MCC, depends on appropriate and adequate documentation. The *coding* of a DRG for an inpatient admission derives from the patient's medical chart/record, which includes (1) admission notes, (2) patient progress notes during a hospital stay, and (3) the discharge summary. The EMR and in particular the discharge summary is intended to capture the primary and secondary diagnoses that are active. The primary diagnosis determines the base DRG. There is no maximal number of secondary diagnoses that can be entered, but typically only 25 diagnoses are used by CMS and/or grouping software for purposes of generating claims. The discharge summary is also used as a concise way for providers seeing the patient for subsequent care/follow-up to understand what transpired in the inpatient setting.[13]

Medical coding is a process of translating the information in the chart into standardized, billable codes that are recorded in a claim which healthcare providers submit to insurers for payments. In particular, the medical codes in the claim constitute important sources to determine the DRG, that is, the reimbursement for the inpatient stay. CMS also requires documentation in the chart (whether paper or electronic) to substantiate each billed CC or MCC. The criteria specified by CMS typically include a combination of results from the patient history, physical examination, laboratory tests, medical imaging, specialty consultations, hospital course, and more. As a result, compliance with Medicare billing criteria requires substantial efforts and time from healthcare providers.

The following processes could lead to scrutiny by auditors: (1) selecting a base DRG with a higher weight than justified; (2) coding a CC or MCC modifying the base diagnosis that is not present/not sufficiently documented; (3) services that are not medically justified by the clinical situation; (4) coding selected diagnoses as present on admission when they were not, and (5) unbundling services/procedures that are bundled under a single DRG.

The adoption of EMRs has substantially changed the way how hospitals bill services. In this paper, we focus on computerized physician order entry (CPOE), an advanced component that is

[13]It is particularly valuable for reconciling medication lists, comparing those medications being used by the individual prior to hospitalization with those prescribed on discharge. Typically they are different, but it is difficult for even the most informed patients to understand what has changed, let alone what needs to be discontinued when a new medication for the same indication is prescribed. The provider seeing the patient on follow-up uses the discharge summary to reconcile the differences.

described as the most typical application in the digitization of medicine because of its substantial impacts on clinician culture and workflow.[14] It is a more advanced type of electronic prescribing that can link to the adverse drug event system to avoid potential medication errors.

An EMR system will typically record the hospital course, providing templates to aid the physician in documentation. For instance, at the time of admitting a patient, the admitting physician can enter a new diagnosis by selecting one from a pop-up list organized by organ system or functional abnormality or simply choose from a populating list of pre-existing diagnoses if the patient has previously been seen in the system. On the one hand, the presence of pop-up lists, preloaded templates, and autofill functions reduces the probability of missing information during the hospital course and makes generating complete medical records much easier. But on the other hand, these functionalities may facilitate inappropriate billing practices, such as enabling "cutting and pasting" or "cloning" information that includes diagnoses and patient status from one note to another for a given patient, despite changes in these parameters in the interim. This process is typically used to save time, but can lead to overpayment or to underpayment, depending on which information is incorrect or incomplete. If a diagnosis that is no longer active is carried forward in cloned information, this could result in overpayment.[15] If a new active diagnosis is not included in the cloned information, this could result in underpayment.

[14]See https://psnet.ahrq.gov/primer/computerized-provider-order-entry.

[15]This leads to upcoding. There are also discussions on EMRs leading to upcoding in outpatient care. See http://www.nytimes.com/2012/09/22/business/medicare-billing-rises-at-hospitals-with-electronic-records.html?_r=0.

2.2 MEDICARE RECOVERY AUDIT PROGRAM

The CMS uses a variety of auditing strategies to ensure the accuracy of Medicare payments, including the Comprehensive Error Rate Testing (CERT) program, Recovery Audit Program (RAP), Medicare Administrative Contractors, Supplemental Medical Review Contractors, and the Zone Program Integrity Contractor audits.[16] As one of the contractors to ensure providers follow Medicare reimbursement policy, the primary task of a recovery auditor is to examine paid claims under Medicare Parts A and B and identify whether the claim contains any improper payments (overpayments or underpayments). The RAP started as a demonstration in six states between March 2005 and March 2008, and was expanded nationwide by 2010. It is unique and distinct from other programs because of its ability to conduct widespread post-payment review.

The CMS contracts the recovery audit work with four independent contractors, each of which is responsible for claims in a geographically defined region that is about one-quarter of the country.[17] RAs perform three types of review: automated, semi-automated, and complex. The first two approaches are mainly based on claims data analysis, with the latter sometimes requiring providers to submit supporting documents for substantiation. Complex reviews must be conducted by a qual-

[16] Zone Program Integrity Contractors were formerly known as the Program Safeguard Contractors. See the Medicare Program Integrity Manual (https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Internet-Only-Manuals-IOMs-Items/CMS019033). Beginning in Fiscal Year (FY) 2016, CMS developed the Unified Program Integrity Contractor strategy to consolidate the integrity efforts of many of the above programs. See https://www.lilesparker.com/unified-program-integrity-contractor-upic-audits-investigations/. The fiscal year is the accounting period for the federal government, from the fourth quarter of the previous year to the third quarter of the current year.

[17] Appendix Figure A1 displays the map of the four RAP regions.

ified healthcare coder or clinician, who must review supporting medical records before making the determination. While most of the reviews are performed in an automated way, on average, over 88 percent of the improper payments are identified from complex review. When an improper payment is identified, the involved provider will receive a notification letter with the rational for the determination. After that, the provider either fulfills the payments or initiates a discussion or an appeal process.[18]

RAs are paid based on a contingency fee basis, a percentage of the corrected payments, ranging from 9.0 – 12.5 percent for all claim types except for claims on durable medical equipment.[19] Note that RAs will get paid in *both* cases of overpayments and underpayments, but need to return the fee if the determination is overturned at all levels of appeal. As a result, RAs are incentivized to accurately identify improper payments, which is confirmed by the recovery audit validation contractor who reviews a monthly random sample of claims that RAs adjudicated as improperly paid and finds that the assessment of RAs are accurate in over 90% of the cases.

The majority of improper payments are overpayments, and most of the corrected payments—81% of the overpayments and 76% of the underpayments—arise from inpatient claims. Thus, the billing and documentation in inpatient hospital care plays an important role in ensuring the compliance with Medicare's payment criteria and documentation and billing requirements, which is the focus of the discussion in this paper. According to the RAP reports, the most common reasons for improper payments are related to the services (1) for which the documentation submitted does not support the services rendered; or (2) that do not meet Medicare's coverage and medical necessity

[18]In case of an underpayment, the provider will be notified in the letter about the repayment process.

[19]The rate went up to 10 – 14.4 percent in FY 2016 after the procurement and contract modification.

criteria; or (3) that are incorrectly coded (Recovery Audit Program, 2011 – 2014).[20] The results of the RAP help CMS identify program vulnerabilities and point out directions on future corrective actions that can be implemented to reduce improper payments.[21]

3 Data and Summary Statistics

The first dataset comes from the audit-level administrative data on the RAP, from which we derive the amount of overpayments and underpayments. The dataset covers 100% audits in the program, including all claims in Medicare Parts A and B that have been reviewed by RAs. We extract the data on hospital inpatient care that extends from 2010 to 2019, with approximately 450 million audits of inpatient admissions.[22] The data include claim characteristics (such as the service dates, service providers, diagnosis codes, procedure codes, DRGs, original paid amounts, and so on) and information on the audits (including the date selected for review, review decision, corrected amounts, and appeal outcomes). We obtain information on inpatient discharges by provider and service from the Inpatient Utilization and Payment Public Use File from the CMS.[23] With these variables, we

[20] These are the key areas in proper billing and documentation that CMS emphasizes and other audit contractors (such as the CERT program) also focus on.

[21] Vulnerability is defined as a claim type or series of related claim types that are more susceptible to improper payments and thus impose greater financial risk to the Medicare program (Recovery Audit Program, 2012).

[22] The dataset also contains claims in other settings, such as skilled nursing facilities, home health, hospice, outpatient care, and durable medical equipment. Also, to focus on the claims paid on IPPS, we exclude claims from Maryland and critical access hospitals in the analysis.

[23] See https://data.cms.gov/provider-summary-by-type-of-service/medicare-inpatient-hospitals/medicare-inpatient-hospitals-by-provider-and-service.

construct the outcome measures in the analysis — the average overpayments / underpayments per discharge at the hospital-DRG level.

We use the Healthcare Information and Management Systems Society Analytics Database to construct the health IT adoption variable. The database is an annual survey, recording the demographic and automation information of the majority of U.S. hospitals and evaluating 90 software applications and technologies. Specifically, it includes information on a hospital's adoption status, year of adoption, component installed, and the identify of the vendor. Following prior studies, we define a hospital to have adopted EMRs if CPOE is live and operational within the organization (Lee et al., 2013; McCullough et al., 2016; Ganju et al., 2021).[24] Figure 1 displays the average adoption rate over time. At the beginning of the sample period, about 33.9% of hospitals adopted CPOE, and the fraction went up to over 98% in 2019.

The health IT database also includes key demographic information about hospitals, such as the number of beds and location. We complement the adoption data with the American Hospital Association Annual Survey (which is available for a shorter time period between 2005 and 2010), using the Medicare provider number and geographic information to link the datasets. Thus, we include a rich set of hospital controls, including bed size, profit status, whether a hospital is a teaching hospital, total admissions, and percentage Medicare and Medicaid discharges. We link the adoption data with audit data using the National Provider Identifier and obtain approximately 3000 hospitals per year, covering about 70% of the hospitals enrolling in Medicare.

Table 1 presents the summary statistics for the main variables by adoption status. The average

[24] We use an alternative adoption definition in the robustness check. The main findings hold. Section 5 provides more details.

overpayment (underpayment) per discharge is $354 ($19), and the amount of overpayments (underpayments) is lower (higher) among adopters. Also, EMR hospitals tend to be larger in terms of bed size, the number of discharges per DRG, or total admissions. They are also more likely to be not-for-profit hospitals or teaching hospitals.

Finally, we merge our data with information on DRG weights from the CMS. We control the weight of a particular DRG, which reflects the average resources used to treat patients in that DRG and could affect hospital billing/coding practices. Table 2 shows the statistics for the DRGs in our empirical analysis. The number of DRGs varies by year, because the discharge data are available only for the top 70 DRGs in 2010 and the top 100 DRGs in years between 2011 and 2013. Though we obtain the discharge data for all DRGs beginning in 2014, some of them are missing after we merge all the data sources. The mean DRG weight changes over time, with substantial variations across DRGs, as reflected in the standard deviations.

4 Empirical Strategy

4.1 MAIN SPECIFICATION

Our empirical specification is motivated by the two potential mechanisms for how EMRs affect hospital documenting and billing services: the *bill-inflation* mechanism and the *complete-coding* mechanism.[25] These two mechanisms lead to similar patterns in claims data—higher bills among EMR hospitals, but result in different audit findings. If the former is the primary mechanism, RAs

[25] EMRs could affect billing and documentation in other parts of care in Medicare. Our discussion focuses on inpatient hospital care, considering that it is the largest contributor to Medicare spending and improper payments.

are likely to make an improper payment determination after a complex review, because patients' conditions could be mis-reported or inadequately justified. However, if the complete-coding mechanism plays a major role, the original claim determination will probably not change after RAs' examination, because the (complete) documentation should comply with Medicare guidelines and support the level of services billed.

We examine how the collected overpayments / underpayments vary as a hospital adopts EMRs. For each year, we first add up the improper payments to the hospital-DRG level and then divide the sum by the number of inpatient discharges at the same level. We estimate the regression equation in the following form:

$$\ln(Y^M_{idt}+1) = \beta^M \text{Adopt}_{it-3} + \alpha^M \text{Weight}_{dt-3} + \gamma^M X_{i,2007} \times t + \delta^M_{id} + \mu^M_{rt} + \varepsilon^M_{idt}, \tag{1}$$

where Y^M_{idt} denotes the average improper payments (in dollars) per inpatient discharge among patients reported with DRG d at hospital i in year t when the audit occurs. Note that RAs are allowed to review claims that were paid within the past three years, and thus, Y^M_{idt} is a ratio of the amount of improper payments between $t-3$ and t to the total number of inpatient discharges during the same period. $M = OP$ (UP) if we consider overpayments (underpayments). We take the natural log of the dependent variable to reduce the skewness in the payment data and for ease of interpretation.[26] We use $\log(Y^M_{idt}+1)$ because the improper payments can be zero, making $\log(Y^M_{idt})$ undefined.[27]

[26] Appendix Figure A2 displays the distribution of the average overpayments and underpayments in the sample, respectively. Both variables have a long tail towards the high values, representing that the majority of improper payments are relatively small.

[27] Brown (2019) transforms the dependent variable in a similar way to avoid the decrease in sample size.

Adopt_{it-3} represents hospital i's adoption status three years ago. Since the outcome measure involves the improper payments in the last three years, we use the three-year lag of adoption status to capture the contemporaneous and subsequent effects on hospitals' billing practices. Similarly, we include the three year lag of the DRG weights.[28]

We also include a set of hospital controls, denoted as $X_{i,2007}$ for which we take the 2007 values, and interact them with a linear time trend. We emphasize the 2007 baseline to also avoid potential changes in these characteristics due to EMR adoption, sharing a similar idea to that in Dranove et al. (2014).[29] δ_{id}^{M} denotes hospital-DRG fixed effects, to capture any time-invariant factors at this level that may simultaneously affect the audit outcomes and the adoption of health IT, such as the persistent patterns of billing practices for a particular DRG at a given hospital. μ_{rt}^{M} is the auditor-/region-year fixed effects, allowing for unrestricted, differential trends by RAs/regions to capture time-varying unobservables in auditing strategies or region characteristics in terms of patient population, medical practice patterns, and so on. For instance, CMS prohibited RAs to review inpatient hospital patient status between FY 2014 and FY 2015, and μ_{rt}^{M} could capture how different RAs respond to this policy change.[30]

The key variable of interest is the adoption status, and the associated coefficient, β^{M}, measures

[28]Note that, on average, the distribution of reviewed claims across t, $t-1$, $t-2$, and $t-3$ is 24%, 36%, 27%, and 13%, respectively. Given the relatively even distribution, we use the three-year lag for these variables.

[29]We use the year 2007 as the baseline for two reasons. First, it is prior to the national rollout of RAP. Second, in our analysis, the earliest year for the claims reviewed by RAs were paid in 2008. In our robustness checks, we also use the year 2002 and 2005 as the baseline year, respectively. The Medicare Modernization Act was enacted in 2003, which required CMS to demonstrate the applicability of RAs to identify improper payments, whereas the Tax Relief and Healthcare Act of 2006 allowed the RAP to be extended to all states (Page 57808, CMS and HHS (2011)).

[30]Note that our specification implicitly includes year fixed effects, which are perfectly collinear with μ_{rt}^{M}.

the effect of adoption on the amount of improper payments. In the case of overpayments, we anticipate β^{OP} to be positive if EMRs mainly facilitate upcoding, which can be captured in the complex review by auditors. However, β^{OP} is expected to be insignificant if the effect on complete documentation dominates. In this latter case, EMRs help capture charges for the services provided by producing supporting documentation adequate to pass audits. In the case of underpayments, we expect β^{UP} to be negative in both mechanisms. For the bill inflation mechanism, hospitals have no desire for down-coding in response to financial incentives. For the complete coding hypothesis, EMRs assist healthcare providers in billing the appropriate level of services, decreasing down-coding. In all, the estimated β's help shed light on which mechanism dominates: the presence of a positive β^{OP} and a negative β^{UP} supports the bill-inflation mechanism, whereas the combination of an insignificant β^{OP} and a negative β^{UP} are in favor of the complete-coding mechanism.

4.2 DISCUSSION ON IDENTIFICATION

Our estimation strategy relies on a fixed effects approach, and thus, the identification is based on variations within a hospital/DRG across time. The key identifying assumption is that improper payments of adopters and non-adopters would have trended similarly in the absence of adoption. Given that EMR and non-EMR hospitals could be different in many aspects, as shown in Table 1, we include a rich set of hospital characteristics and hospital fixed effects in the estimation. However, concerns regarding omitted variable could still exist, as there might be time-varying unobservables that are correlated with hospital adoption decision and billing practices at the same time. For instance, hospitals with more experience in using computers and specialized software could adopt EMRs earlier and also have better capabilities in managing financial transactions.

To investigate whether EMR hospitals share a common trend with non-EMR hospitals prior to adoption, we estimate Equation (1) but replace the adoption variable with an indicator variable for eventual adopters—the treatment group—interacted with a set of dummies for each year before and after adoption. Figure 2 presents the coefficients on the interactions between EMR adoption and time since adoption, as well as their 95% confidence intervals. All reported coefficients are relative to $t = -1$, one year before adoption. Visually, there were no differential pre-adoption trends for EMR hospitals relative to non-EMR hospitals in both types of improper payments. The last row of Table 3 shows the p-value of the test for the joint insignificance of the coefficients for the pre-adoption periods. We cannot reject the null hypothesis that the pre-adoption trend of collected overpayments/underpayments among EMR hospitals is not statistically different from that in non-EMR hospitals.

Despite no differential pre-adoption trends, one might still worry about trend breaks in omitted variables that occur at the same time as EMR adoption and could also lead to parallel pre trends. For instance, certain types of hospitals might change the hiring of coders or administrative staff post adoption, leading to different trends in administrative costs before and after EMR adoption. Or, they might shift their payer composition via changing patient composition to reduce the financial burdens from adopting EMRs. To address this concern, we examine other outcome measures, such as administrative costs and fraction of Medicare discharges, to see whether there were differential changes in hospital administrative staffing or payer split around the time of adoption. We measure the former using the administrative and general costs, which include a wide variety of provider administrative costs on administrative services, legal and accounting services, human resources, and so on. We calculate the latter as the ratio of number of Medicare discharges to total discharges for a given hospital. All these data come from the the Healthcare Cost Report Information System. We use an event study design similar to the one from which we obtain the estimates for Figure 2 and plot the interaction coefficients and their 95% confidence intervals in Appendix Figure A3. In both cases, there is no trend in either the pre- or post-adoption periods.[31]

Recent literature has raised concerns about interpreting the estimates from the two-way fixed effects regression as causal effects when the treatment timing is staggered and the treatment effect varies over time within the same unit or between groups that get treated at different periods (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021). In our setting, hospitals adopt EMRs in different years during the sample period, leading to a staggered treatment timing. To examine whether the results from our two-way fixed effects estimator are severely biased, we re-estimate Equation (1) following De Chaisemartin and d'Haultfoeuille (2020) and De Chaisemartin and D'Haultfoeuille (2022). The general idea is to compare the $t-1$ to t improper payment changes of hospitals going from without EMRs to with EMRs from $t-1$ to t, and of hospitals that remain with no EMRs at both periods. The findings on the pre-adoption trends and the estimated treatment effects are generally consistent, as shown in Columns (1) and (2) of Appendix Table A3.

[31]The pre-adoption indicators are not jointly statistically significant for both outcome measures (p-value=0.227 for administrative costs and p-value=0.132 for the fraction of Medicare discharges). The post-adoption coefficients are statistically insignificant for the fraction of Medicare discharges (p-value=0.436), but jointly statistically significant for the administrative cost (p-value=0.026), which might simply reflects both positive and negative values in different years.

4.3 HETEROGENEITY ANALYSIS

To further explore the underlying mechanism, we estimate the heterogenous effects of EMR adoption based on the capabilities of RAs. As pointed out by Ganju et al. (2021), certain RAs have developed specialized capabilities of identifying inappropriate documentation in EMRs and thus are more effective in capturing improper payments. To verify this implication, we estimate the following equation:

$$\ln(Y^M_{idt}+1) = \beta^M_1 \text{Special}_r \times \text{Adopt}_{it-3} + \beta^M_2 \text{NotSpecial}_r \times \text{Adopt}_{it-3} \tag{2}$$

$$+ \alpha^M \text{Weight}_{dt-3} + \gamma^M X_{i,2007} \times t + \delta^M_{id} + \mu^M_{rt} + \varepsilon^M_{idt},$$

where $\text{Special}_s = 1$ ($\text{NotSpecial}_s = 1$) if hospital i is assigned to auditor r that possesses (lacks) the specialized capabilities.[32] The values of β^M_1 reflect the extent to which such capabilities help identifying the improper payments arising due to EMR adoption. If an EMR is more of a tool to game the reimbursement system, especially due to the functionalities of pre-loaded templates and "cloned" records, we expect $\beta^{OP}_1 > \beta^{OP}_2 > 0$. Other variables share the same definitions in Equation (1).

Moreover, the impact from EMRs may vary by hospital type. Prior studies have found that for-profit hospitals or financially distressed hospitals are more likely to upcode patients due to financial incentives (Silverman and Skinner, 2004; Dafny, 2005; Li, 2014). While for-profit hospitals demonstrate greater incentives for upcoding, not-for-profit hospitals are more likely to pick up advanced EMR systems earlier, as suggested by Adler-Milstein et al. (2017) and as shown in our

[32]We thank Ganju et al. for providing us the information on the auditors with the specialized capabilities.

data (Table 1). As a result, it remains unclear in which type of hospitals the impact would be more significant. To examine this, we re-estimate Equation (1) separately for each type of hospitals.

Several points are worth noting. First, the overpayments or underpayments we examine might not represent the prevalence of improper payments. Both of them are conditional on the billing error being substantial enough that RAs want to report. Consider an optimization problem facing auditors who attempt to maximize revenues subject to time/resource constraints. RAs are probably most interested in claims with obvious billing errors or resulting in a large amount of corrected payments. It is likely that some errors are not reported because they are not either sufficiently egregious or sufficiently lucrative. However, we believe that RAs also have incentives to ensure the accuracy of their audits because they will not get reimbursed if the claim is overturned in appeal. Moreover, we believe that our measures of improper payment incidence might not be too far away from the actual prevalence of improper payments. Appendix Table A1 shows the overall improper payment rates due to errors other than medical necessity in the RAP and CERT program. Note that CERT auditors are distinct from other auditors in the sense that they review a *random* sample of Medicare Parts A and B claims.[33] The difference in the overall improper payment rate between both programs is not substantial especially after 2012. We acknowledge that our outcome

[33] An arguably better outcome would be the improper payments determined by the CERT auditors. However, very limited information on CERT audits is publicly available. The overall improper payment rate in RAP is a ratio of the total recovered payments to the total payments in the reviewed claims, whereas the improper payment rate in CERT is obtained from the Medicare Fee-for-services Improper Payments Reports (Supplementary Appendices, CERT Program, 2021). We exclude errors associated with medical necessity from the calculation, because RAs became aggressive in reviewing inpatient claims for medical necessity between 2011 and 2013, which resulted in a large number of appeals.

measure might not be representative enough, but our results inform the directional changes in the occurrence of (obvious) upcoding errors as a hospital adopts EMRs.

Second, we examine the extent to which EMRs affect coding decisions rather than treatment decisions. Hence, we do not consider cases such as EMRs leading to an additional test to be ordered that would not have been done without the technology, which could occur if, for instance, EMRs remind the physician to perform a diagnostic test. Instead, our analysis is conditional on all the services that have been documented in the medical record and examines how the application of EMRs affects the tasks of capturing the care provided by the hospital which can be translated into bills or claims, through the process of coding. The EMR may change what is coded, and in that sense increase the reimbursement, but the purpose of our analysis is to examine whether that change is justified as determined by audit.

5 Results

5.1 MAIN RESULTS

Table 3 presents the estimated coefficients for the key variables of interest, with the the first two columns for Equation (1) and the last two for Equation (2). Because we take the natural log of the dependent variable, the coefficients inform the percent change in improper payments as a result of adoption. The effect on overpayments is insignificant, whereas EMR adoption decreases underpayments by 7.06 percent as shown in Column (2). Moreover, the effect of EMR adoption on lowering underpayments seems to get larger over time, as shown in Figure 2. The last row shows the *p*-values for testing whether there were differential pre-adoption trends in improper payments for eventual adopters relative to non-adopters. We cannot reject the null hypothesis for both outcome measures, suggesting that the trends of identified improper payments among EMR hospitals prior to their adoption are not statistically different from those among non-EMR hospitals.

The results for overpayments are similar after taking into account whether the RA is capable of identifying the use of over-documentation and copied records in EMRs, as shown in Column (3) of Table 3. In the case of underpayments, however, the coefficient on the interaction between adoption and the indicator for RAs with the specialized capabilities is significantly negative, suggesting that EMR hospitals located in states assigned to these RAs experience a smaller amount of underpayments relative to non-EMR hospitals; that is, hospitals with EMRs tend to make relatively fewer "mistakes" that lead to underpayments, which are identified by auditors with the superior auditing capabilities. The estimate corresponds to an average decrease of 7.37 percent in underpayments as a consequence of adoption. In other words, the significantly negative effect observed in the overall sample is mainly driven by the auditors with specialized capabilities. While these capabilities mainly target over-billing practices, RAs with expertise probably went through special training in detecting inappropriate documentation that could lead to both overpayments and underpayments, and thus developed better capabilities to identify such instances.

Taken together, the finding of insignificant effects of EMR adoption on overpayments, regardless of the RAs' capabilities, along with EMRs lowering underpayments, provides supporting evidence for the complete-coding mechanism. It suggests that EMRs improve data collection and documentation during the hospital course and enhance allowable billing. These results are consistent with the findings in prior studies—even those concluding more complicated conditions reported as a result of EMR adoption—in the sense that EMR hospitals are able to bill the appropriate level instead of a lower billing code, by using more complete documentation. As a result, RAs especially those with specialized capabilities identify a disproportionately lower amount of underpayments in hospitals with EMRs.

5.2 AUDIT INITIATIONS BY EMR ADOPTION

EMR adoption could affect the cost to recover payments from a claim. On the one hand, RAs may face higher costs to examine claims from providers with EMRs, because the technology assists in documentation of care and could produce a lot of extra files that require RAs more efforts. In this case, RAs would be less likely to review claims from hospitals with EMRs, which could lead to the null effect on overpayments as found above, even though EMR hospitals do engage in over-billing practices. But on the other hand, the standardized data formats generated from EMRs may facilitate the inspection of claims, considering that the use of statistics and data mining has become commonplace in identifying instances of improper billing. Moreover, the Office of Inspector General issued a report describing how EMRs could potentially result in billing fraud.[34] Claims from EMR hospitals may be subject to heightened scrutiny, as RAs may anticipate more erroneous billing identified from these claims. As a result, it remains uncertain to what extent a provider's adoption status affects whether his/her submitted claims get selected for review. To answer this question and examine whether the main results are driven by the difference in audit initiation by EMR adoption, we replace the dependent variable with audit rates—the average number of audits per discharge. Appendix Table A4 suggests that whether a hospitals has adopted EMRs makes little difference in initiating audits.

[34] See https://oig.hhs.gov/oei/reports/oei-01-11-00571.pdf.

5.3 RESULTS BY HOSPITAL TYPE

To investigate whether the effect of EMRs varies by hospital type, we re-estimate the main specification separately for different subsamples. Prior literature has found that for-profit hospitals have greater incentives to engage in upcoding than not-for-profit hospitals (Silverman and Skinner, 2004; Dafny, 2005; Li, 2014). We first examine this hypothesis by rerunning Equation (1) for for-profit and not-for-profit hospitals, respectively. Columns (1) and (2) in Table 4 present the coefficient for the key variable of interest for not-for-profit hospitals. We do not see any significant impact from EMRs on overpayments but a decrease of 7.1 percent in underpayments as a result of EMR adoption among these hospitals. When it comes to for-profit hospitals, EMRs have little impact on both overpayments and underpayments. Again, we find no evidence supporting the bill-inflation mechanism among for-profit or not-for-profit hospitals. In fact, the finding that EMR hospitals experience relatively fewer underpayments mainly occur to not-for-profit hospitals.

Next, we examine whether EMR adoption shows differential effects on financially distressed hospitals and financially non-distressed hospitals. A hospital is defined to be the former (latter) if its debt-asset ratio is above (below) the median. Columns (5) and (6) in Table 4 present the estimates for the former. Again, the coefficient for EMR adoption remains statistically insignificant in overpayments, but becomes significantly negative in underpayments. On average, financially distressed hospitals with EMRs experience a decrease of 10.6% in underpayments relative to those without EMRs, suggesting that EMRs could play an important role in recovering billable services for these hospitals. The last two columns report the results for hospitals that undergo less financial distress and show no impact of EMR adoption on improper payments. Taken together, we still find no evidence for the bill-inflation mechanism, but the results in the heterogeneity analysis continue lending support to the complete-coding mechanism.

5.4 ROBUSTNESS CHECK

We test the sensitivity of the results in the following ways. First, we replace the total improper payments with those more related to upcoding. As mentioned in Section 2.2, improper payments mainly occur to services that lack adequate supporting documentation, are incorrectly coded, or fail to meet the Medicare's coverage and medical necessity criteria. Upcoding could be associated with any of them, but we view the first two the most relevant. In both cases, the improper payments arise due to preferential selection of billing codes leading to higher reimbursement than is supported by medical record documentation. In the context of hospitalization care, billing errors related to medical necessity are more about where the services should be provided—in the inpatient or ambulatory setting (CERT Program, 2013), but the services billed could be correctly coded.[35] The audit data contain information on why a claim is selected for review. To construct the upcoding-related improper payments, we remove all the claims that are audited due to medical necessity. Appendix Table A2 reports the results. The effect of EMRs is insignificant on overpayments but significantly negative on underpayments, with an even larger magnitude than that in the overall sample.

Second, we use an alternative adoption definition—a hospital is defined to have adopted EMRs

[35]Medically unnecessary care could also be a broader issue, touching areas such as pharmaceuticals, home health care, hospice care, and so on.

if CPOE or the component, physician documentation, is live and operational in the hospital.[36] Physician documentation is also an advanced application, offering physicians structured templates to document a patient's daily progress, operative notes, consultation notes, emergency department visits, discharge summary, and other relevant information during the course of a hospital admission. The last two columns in Appendix Table A3 report the results. The main findings hold: EMR adoption has little impact on overpayments but decreases underpayments by almost 8 percent, yet with marginal significance. Third, we rerun regression Equation (1) using a recently-developed difference-in-difference estimation approach (De Chaisemartin and d'Haultfoeuille, 2020; De Chaisemartin and D'Haultfoeuille, 2022), given the concerns on the potential bias in the estimates from the two-way fixed effects model. The results are robust, as shown in the first two columns of Appendix Table A3.

Finally, we assess the robustness of the main results using different specifications. Appendix Table A5 reports the results across six specifications for overpayments and underpayments, respectively, by progressively adding more controls or fixed effects. The estimates get closer to those from our preferred specifications (in Column (4)) after hospital-DRG fixed effects are included. Considering that hospitals may be aware of the implementation of RAP before 2007, we also try 2002 and 2005 as the baseline year, respectively, and present the results in the last two columns of Appendix Table A5.

[36]Another component in the EMR system, clinical data repository, could also be relevant to documentation of care (Dranove et al., 2014). Unlike CPOE or physician documentation, it is a basic component. Over 87% of hospitals had adopted clinical data repository by 2010. Given the limited variation in the adoption of clinical data repository during our sample period, we focus on CPOE or physician documentation in our paper. Thus, our results imply the effect from the advanced adoption of EMRs rather than the basic adoption.

6 Conclusion

We study the correlation between the amount of recovered improper payments and hospital EMR adoption. Using audit data allows us to distinguish between the two potential mechanisms: the bill-inflation mechanism and the complete-coding mechanism, which cannot be separated in conventional claims data. Since inpatient hospital care accounts for the largest proportion in Medicare spending, the findings here shed light on the underlying effect of EMR adoption on hospital coding behavior. We find no evidence supporting the proposition that EMRs facilitate bill inflation, but more evidence pointing to the complete-coding mechanism during our sample period.

Our lack of substantiation to the bill-inflation mechanism suggests that more focus could be concentrated on other common types of billing errors, such as the improper payment error of lacking medical necessity or the improper billing in Medicare Part B that is billed under fee-for-services. Instead, compliance with Medicare billing criteria could be costly, and the application of EMRs makes the process less complicated. Policy makers might consider providing more guidance to providers and create the right incentive for them to engage in proper documentation.

References

Adler-Milstein, J., Holmgren, A. J., Kralovec, P., Worzala, C., Searcy, T., and Patel, V. (2017). Electronic health record adoption in us hospitals: the emergence of a digital "advanced use" divide. *Journal of the American Medical Informatics Association*, 24(6):1142–1148.

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

Barros, P. and Braun, G. (2017). Upcoding in a national health service: the evidence from Portugal. *Health economics*, 26(5):600–618.

Bauder, R., Khoshgoftaar, T. M., and Seliya, N. (2017). A survey on the state of healthcare upcoding fraud analysis and detection. *Health Services and Outcomes Research Methodology*, 17(1):31–55.

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

Brown, Z. Y. (2019). Equilibrium effects of health care price information. *Review of Economics and Statistics*, 101(4):699–712.

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

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

CERT Program (2011-2021). Medicare fee-for-service improper payments report. *Centers for Medicare and Medicaid Services*.

CMS and HHS (2011). Federal Register. *Federal Register*, 76(80):57625–57896.

Cook, A. and Averett, S. (2020). Do hospitals respond to changing incentive structures? Evidence from Medicare's 2007 DRG restructuring. *Journal of Health Economics*, 73:102319.

Dafny, L. S. (2005). How do hospitals respond to price changes? *American Economic Review*, 95(5):1525–1547.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

De Chaisemartin, C. and D'Haultfoeuille, X. (2022). Difference-in-differences estimators of intertemporal treatment effects. Technical report, National Bureau of Economic Research.

Dranove, D., Forman, C., Goldfarb, A., and Greenstein, S. (2014). The trillion dollar conundrum: Complementarities and health information technology. *American Economic Journal: Economic Policy*, 6(4):239–270.

Fang, H. and Gong, Q. (2017). Detecting potential overbilling in Medicare reimbursement via hours worked. *American Economic Review*, 107(2):562–91.

Ganju, K. K., Atasoy, H., and Pavlou, P. A. (2021). Do electronic health record systems increase medicare reimbursements? The moderating effect of the recovery audit program. *Management Science*.

Geruso, M. and Layton, T. (2020). Upcoding: evidence from Medicare on squishy risk adjustment. *Journal of Political Economy*, 128(3):984–1026.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Gowrisankaran, G., Joiner, K., and Lin, J. (2022). How do hospitals respond to Medicare payment reforms? *NBER Working Paper No. 22873*.

Januleviciute, J., Askildsen, J. E., Kaarboe, O., Siciliani, L., and Sutton, M. (2016). How do hospitals respond to price changes? Evidence from norway. *Health economics*, 25(5):620–636.

Jürges, H. and Köberlein, J. (2015). What explains drg upcoding in neonatology? The roles of financial incentives and infant health. *Journal of health economics*, 43:13–26.

Lee, J., McCullough, J. S., and Town, R. J. (2013). The impact of health information technology on hospital productivity. *The RAND Journal of Economics*, 44(3):545–568.

Li, B. (2014). Cracking the codes: Do electronic medical records facilitate hospital revenue enhancement? *Working paper*.

McCullough, J. S., Parente, S. T., and Town, R. (2016). Health information technology and patient outcomes: The role of information and labor coordination. *The RAND Journal of Economics*, 47(1):207–236.

Qi, K. and Han, S. (2020). Does IT improve revenue management in hospitals? *Journal of the Association for Information Systems*, 21(6):7.

Recovery Audit Program (2011-2016). Recovery auditing in Medicare for Fiscal Year 2011 – 2016. *Centers for Medicare and Medicaid Services*.

Sacarny, A. (2018). Adoption and learning across hospitals: The case of a revenue-generating practice. *Journal of health economics*, 60:142–164.

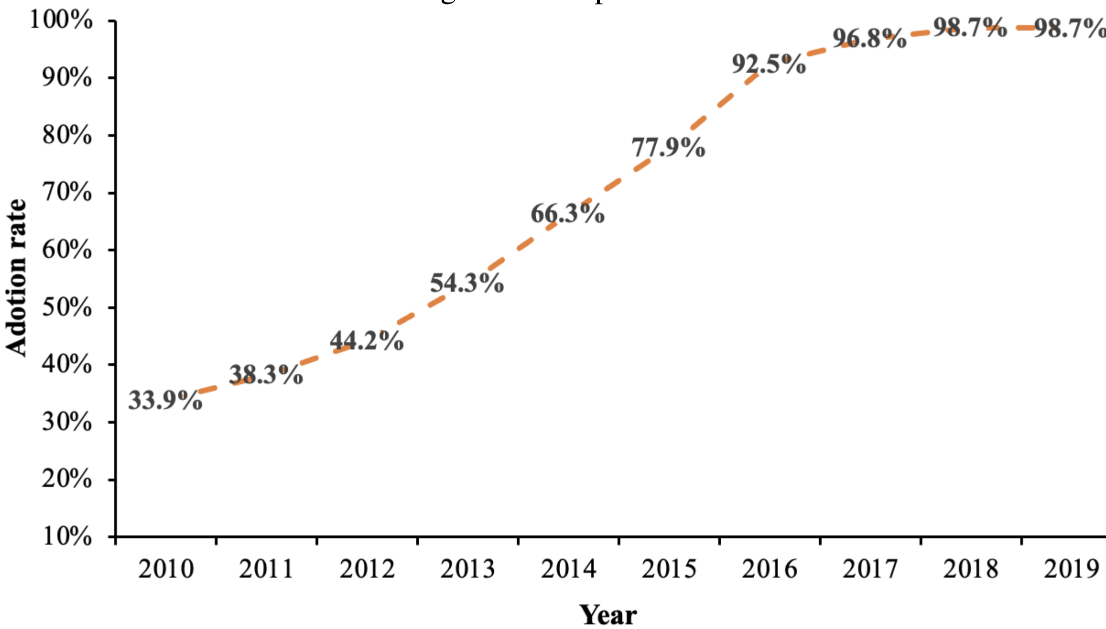
Shi, M. (2021). The costs and benefits of monitoring providers: Evidence from medicare audits. *Working paper*.

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

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

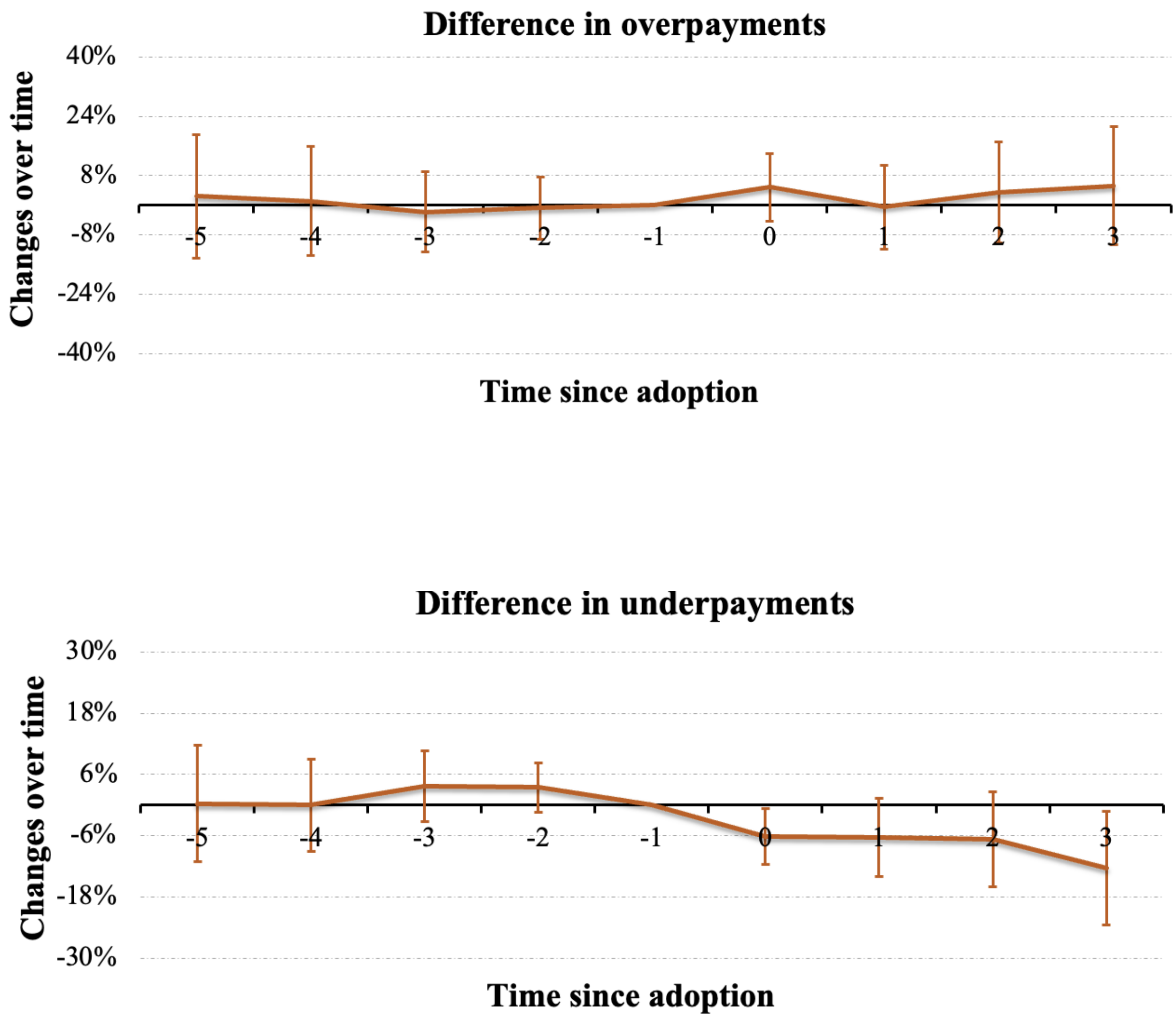
Verzulli, R., Fiorentini, G., Lippi Bruni, M., and Ugolini, C. (2017). Price changes in regulated healthcare markets: Do public hospitals respond and how? *Health economics*, 26(11):1429–1446.

Figure 1: Adoption rate over time



Note: This figure displays the trend of the CPOE adoption rate.

Figure 2: Differences in improper payments over time between adopters and non-adopters



Note: Error bars show 95 percent confidence intervals.

Table 1: Summary statistics for main variables by adoption status

Variable	All hospitals	Non-adopters	Adopters
Overpayments per discharge ($)	354	385	348
	(244)	(312)	(229)
Underpayments per discharge ($)	19	16	20
	(35)	(39)	(35)
Average discharges per DRG	40	31	41
	(29)	(25)	(29)
Staffed beds	225	107	248
	(215)	(105)	(223)
Admissions	10,558	4,171	11,795
	(10,597)	(4,754)	(10,965)
% Medicare and Medicaid discharge	61.90	61.92	61.90
	(13.7)	(16.1)	(13.2)
Teaching hospitals (%)	8.96	1.68	10.4
	(28.6)	(12.9)	(30.5)
Not-for-profit hospitals (%)	63.7	41.8	68.0
	(48.1)	(49.4)	(46.7)
# hospitals	3,022	489	2,533

Note: Table reports the mean values in 2007 and standard deviations in parentheses.

Table 2: Summary statistics for DRGs in the estimation

Year	Mean	Std. Dev.	Min	Max	# DRGs
2010	2.62	1.76	0.540	7.93	70
2011	1.49	1.08	0.550	5.83	100
2012	1.51	1.08	0.543	5.83	100
2013	1.54	1.08	0.554	5.84	100
2014	2.38	2.31	0.551	25.4	351
2015	2.62	2.52	0.601	25.3	281
2016	2.65	2.42	0.674	17.7	181
2017	2.87	2.37	0.657	17.9	166
2018	3.02	3.06	0.696	25.4	141
2019	3.08	2.95	0.653	18.3	87

Note: Information on inpatient discharge per DRG per hospital is only available for the top 70 DRGs in 2010 and top 100 DRGs in 2011–2013. We obtain the discharge data for all DRGs beginning in 2014, but only the number of DRGs reported are used in the estimation after merging all the data sources.

Table 3: Effect of EMR adoption on improper payments

	Overall		By RA capabilities	
	OP	UP	OP	UP
	(1)	(2)	(3)	(4)
Adoption	0.0408	-0.0706***		
	(0.0421)	(0.0272)		
Adoption × RAs **w/** specialized capabilities			0.0521	-0.0737**
			(0.0477)	(0.0305)
Adoption × RAs **w/o** specialized capabilities			-0.0211	-0.0538
			(0.0754)	(0.0552)
N	85,081	85,081	85,081	85,081
P value on joint insignificance of pre-adoption periods	0.975	0.422	–	–

Note: OP (UP) stands for overpayments (underpayments). Other regressors include DRG weights, hospital-DRG fixed effects, RA-year fixed effects, and the following variables valued in 2007 interacted with a linear time trend: number of beds, total admissions, percentage of Medicare and Medicaid discharges, profit status, and whether it is a teaching hospital. We weight each observation by the number of total discharges at the hospital-DRG level. Standard errors in parentheses, clustered at the hospital-DRG level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of EMR adoption on improper payments, by hospital characteristics

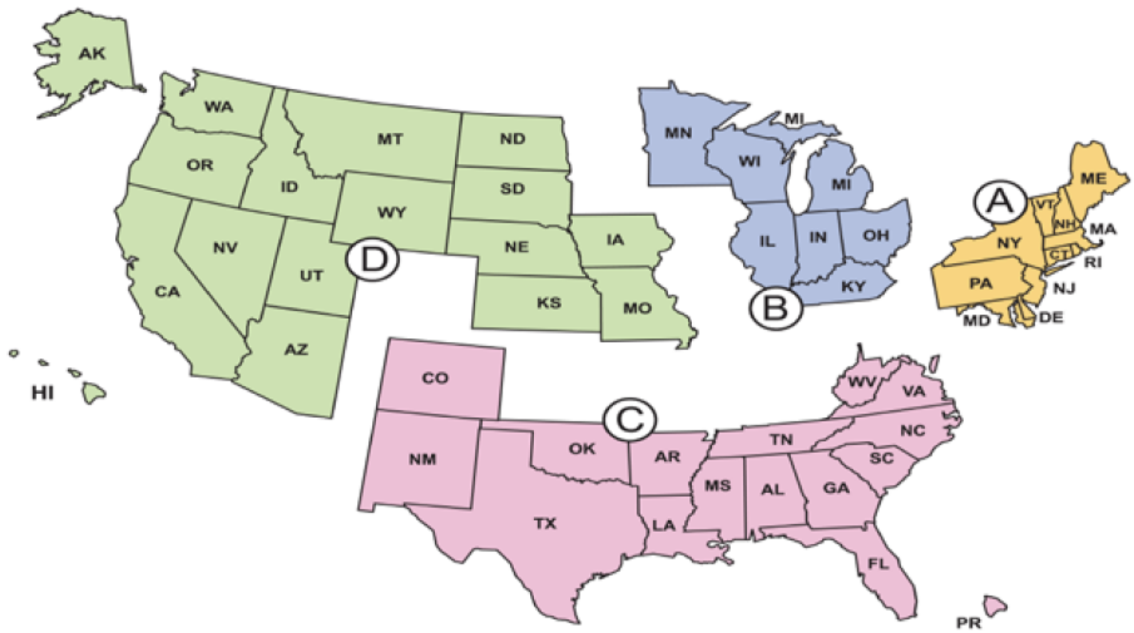
	Not-for-profit		For-profit		Financially distressed		Not financially distressed	
	OP	UP	OP	UP	OP	UP	OP	UP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adoption	-0.00822	-0.0710**	0.117	-0.0344	0.0207	-0.106***	0.0735	-0.0373
	(0.0473)	(0.0298)	(0.103)	(0.0810)	(0.0627)	(0.0370)	(0.0556)	(0.0394)
N	65,670	65,670	15,141	15,141	43,911	43,911	41,170	41170

Note: OP (UP) stands for overpayments (underpayments). Other regressors include DRG weights, hospital-DRG fixed effects, RA-year fixed effects, and the following variables valued in 2007 interacted with a linear time trend: number of beds, total admissions, percentage of Medicare and Medicaid discharges, profit status, and whether it is a teaching hospital. We weight each observation by the number of total discharges at the hospital-DRG level. Standard errors in parentheses, clustered at the hospital-DRG level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix 1: Extra Figures and Tables

Figure A1: Map of RA regions



Note: Below lists the RA for each region: Performant Recovery for Region A, CGI for Region B, Connolly for Region C, and HealthData Insights (HDI) for Region D.

Figure A2: Distribution of average improper payments

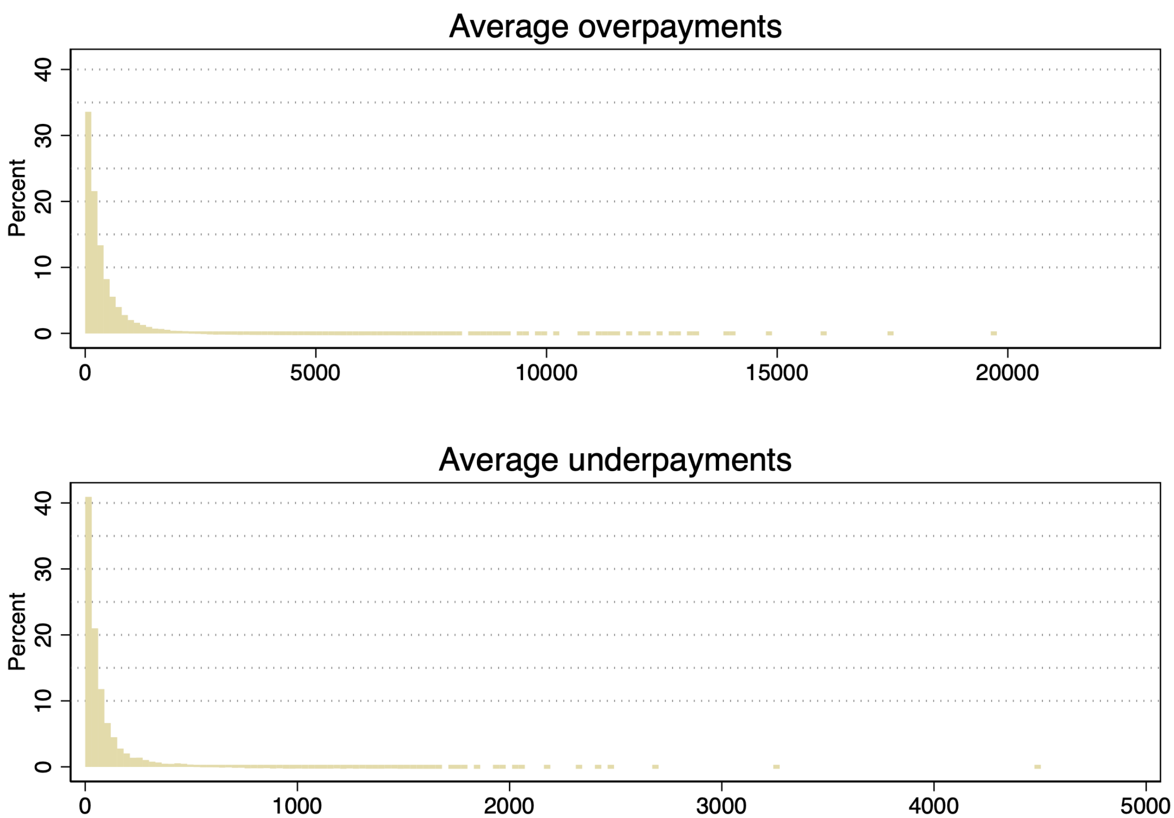
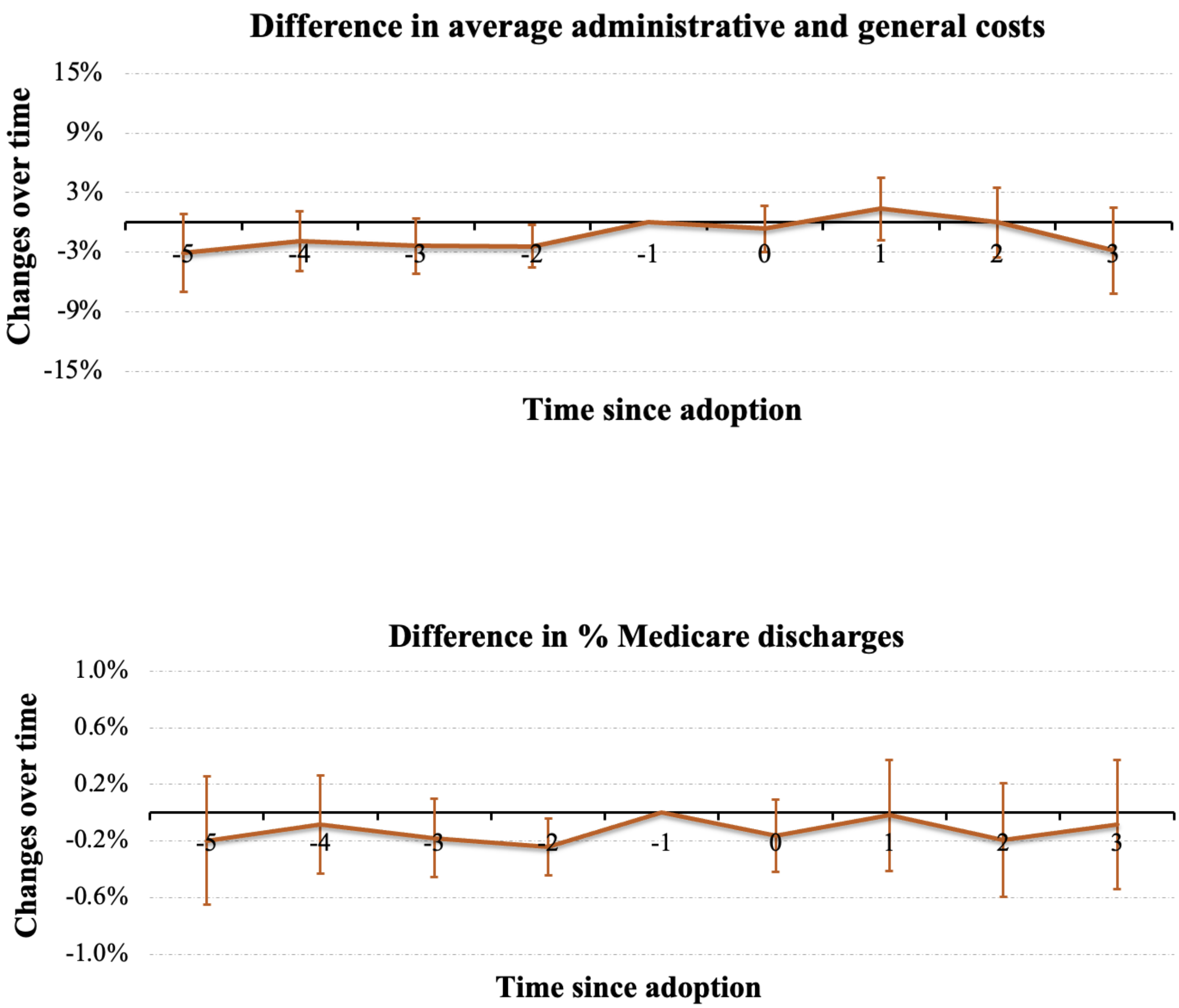


Figure A3: Differences in other outcome measures over time between adopters and non-adopters



Note: Error bars show 95 percent confidence intervals.

Table A1: Overall improper payment rates (%) due to errors other than medical necessity, CERT vs. RAP

Year	CERT	RAP
2011	6.9	10.1
2012	8.9	11.7
2013	11.1	11.3
2014	8.9	7.5
2015	8.7	8.1
2016	6.9	7.7
2017	5.9	5.3
2018	5.9	5.0

Note: The overall improper payment rate in CERT is obtained from Table A6 in the Medicare Fee-for-services Improper Payments Reports (Supplementary Appendices, CERT Program, 2021). Note that we adjusted the improper payment rates in CERT to make them comparable with the improper payment rates in RAP, because starting in 2012, each reporting year in CERT includes claims from July 1^{st} two years before the reporting year to June 30^{th} one year before the reporting year.

Table A2: Effect of EMR adoption on improper payments, excluding medical unnecessary admissions

	Overall		By RA capabilities	
	OP	UP	OP	UP
	(1)	(2)	(3)	(4)
Adoption	0.0883	-0.136***		
	(0.0674)	(0.0456)		
Adoption $\times$ RAs **w/** specialized capabilities			0.114	-0.149***
			(0.0805)	(0.0529)
Adoption $\times$ RAs **w/o** specialized capabilities			-0.0192	-0.0804
			(0.0863)	(0.0803)
N	40,349	40,349	40,349	40,349
P value on joint insignificance of pre-adoption periods	0.105	0.164	–	–

Note: OP (UP) stands for overpayments (underpayments). Other regressors include DRG weights, hospital-DRG fixed effects, RA-year fixed effects, and the following variables valued in 2007 interacted with a linear time trend: number of beds, total admissions, percentage of Medicare and Medicaid discharges, profit status, and whether it is a teaching hospital. We weight each observation by the number of total discharges at the hospital-DRG level. Standard errors in parentheses, clustered at the hospital-DRG level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Effect of EMR adoption on improper payments, using alternative approaches

	Alternative DiD estimation approach		Alternative adopt- -ion definition	
	OP	UP	OP	UP
	(1)	(2)	(3)	(4)
Adoption	0.170	-0.0856**	0.0882	-0.0796*
	(0.130)	(0.0425)	(0.0577)	(0.0417)
N	85,081	85,081	85,081	85,081
P value on joint insignificance of pre-adoption periods	0.210	0.111	0.149	0.930

Note: OP (UP) stands for overpayments (underpayments). Other regressors include DRG weights, hospital-DRG fixed effects, RA-year fixed effects, and the following variables valued in 2007 interacted with a linear time trend: number of beds, total admissions, percentage of Medicare and Medicaid discharges, profit status, and whether it is a teaching hospital. We weight each observation by the number of total discharges at the hospital-DRG level. Standard errors in parentheses, clustered at the hospital-DRG level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Effect of EMR adoption on selecting claims for review

	Overall	By RA capabilities
	(1)	(2)
Adoption	0.00200	
	(0.00221)	
Adoption $\times$ RAs **w/** specialized capabilities		0.00249
		(0.00240)
Adoption $\times$ RAs **w/o** specialized capabilities		-0.000703
		(0.00541)
N	85,081	85,081
P value on joint insignificance of pre-adoption periods	0.0152	–

Note: OP (UP) stands for overpayments (underpayments). Other regressors include DRG weights, hospital-DRG fixed effects, RA-year fixed effects, and the following variables valued in 2007 interacted with a linear time trend: number of beds, total admissions, percentage of Medicare and Medicaid discharges, profit status, and whether it is a teaching hospital. We weight each observation by the number of total discharges at the hospital-DRG level. Standard errors in parentheses, clustered at the hospital-DRG level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effect of EMR adoption on improper payments, in different specifications

	Overpayments					
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption	0.0260	0.0462	0.0531	0.0408	0.0525	0.0449
	(0.0262)	(0.0331)	(0.0429)	(0.0421)	(0.0443)	(0.0425)
Hospital fixed effects	no	yes	yes	yes	yes	yes
Hospital-DRG fixed effects	no	no	yes	yes	yes	yes
Hospital controls valued in year 2007	no	no	no	yes	no	no
Hospital controls valued in year 2002	no	no	no	no	yes	no
Hospital controls valued in year 2005	no	no	no	no	no	yes
N	139,107	139,011	85,538	85,081	78,441	84,343
	Underpayments					
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption	-0.00995	-0.0123	-0.0607**	-0.0706***	-0.0832***	-0.0694**
	(0.0131)	(0.0204)	(0.0271)	(0.0272)	(0.0282)	(0.0272)
Hospital fixed effects	no	yes	yes	yes	yes	yes
Hospital-DRG fixed effects	no	no	yes	yes	yes	yes
Hospital controls valued in year 2007	no	no	no	yes	no	no
Hospital controls valued in year 2002	no	no	no	no	yes	no
Hospital controls valued in year 2005	no	no	no	no	no	yes
N	139,107	139,011	85,538	85,081	78,441	84,343

Note: The four specifications progressively add more controls or fixed effects, as specified in the lower panel. Hospital controls in the last three rows include the following variables valued in the corresponding year interacted with a linear time trend: number of beds, total admissions, percentage of Medicare and Medicaid discharges, profit status, and whether it is a teaching hospital. We weight each observation by the number of total discharges at the hospital-DRG level. Standard errors in parentheses, clustered at the hospital-DRG level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$