Monitoring for Waste: Evidence from Medicare Audits *

Maggie Shi†

September 15, 2022

Abstract

This paper examines the extent to which public programs should monitor for wasteful expenditure. I study a large Medicare program that monitored for unnecessary health-care spending, and consider its effect on government savings, provider compliance costs, and patient health. Every dollar Medicare spent on monitoring generated \$24–29 in government savings. The majority of savings stem from the deterrence of future care, rather than reclaimed payments from prior care. The health of the marginal patient denied care is not harmed, indicating that monitoring primarily deters *unnecessary* care. Instead, the main tradeoff to monitoring is the compliance cost it imposes on providers – for every \$1,000 in Medicare savings, providers incur \$178–218 in higher administrative costs. However, I provide evidence that these costs are driven by the *investments* providers make to improve compliance, like adopting technology to assess the cost-effectiveness of care, rather than the hassle costs of the monitoring process.

*This paper was previously circulated as "The Costs and Benefits of Monitoring Providers: Evidence from Medicare." I am grateful to Wojciech Kopczuk, Adam Sacarny, Pietro Tebaldi, and Michael Best for their input and support. I also thank Jetson Leder-Luis, Jon Skinner, Tal Gross, Ashley Swanson, Bentley MacLeod, Gautam Gowrisankaran, Cailin Slattery, Parker Rogers, Ben Chartock, Kelli Marquardt, Claudia Halbac, Melinda Pitts, Bernard Salanie, Aina Katsikas, Josh Gottlieb, Tim Layton, David Cutler, and other participants at various seminars and conferences for helpful comments and feedback. I thank Mohan Ramanujan, Daniel Feenberg, Elizabeth Adams, Jean Roth, Adrienne Henderson, and Ashley Badami for their assistance in accessing and managing the data. I gratefully acknowledge fellowship support from the Agency for Healthcare Research and Quality (#R36HS027715-01) and the National Institute on Aging (#T32-AG000186). All errors are my own.

†University of Chicago Harris School of Public Policy.

1 Introduction

Much of government spending is contracted out to third parties, leading to concerns about wasteful spending. Contracted goods and services account for 30 percent of federal spending in the U.S., so avoiding waste is a challenging $2 trillion problem (U.S. Government Accountability Office, 2019). Economic theory prescribes a straightforward solution: monitoring third party spending (Laffont and Tirole, 1986; Baron and Besanko, 1987). And yet in practice, monitoring is seemingly underutilized – it is estimated that over half of the losses from wasteful federal spending go undetected (Cunningham et al., 2018). However, if the social cost of monitoring is sufficiently large, then leaving "money on the table" may in fact be the optimal choice.

The question of how much to monitor is particularly salient for healthcare programs like Medicare, the federal insurance program for the elderly and disabled. On the one hand, the sheer magnitude of potential savings makes increasing monitoring an attractive policy tool. All Medicare expenditure is contracted out to healthcare providers, who then have wide latitude over spending decisions. Waste is widespread: up to 13 percent of Medicare spending (two percent of all federal expenditure) goes to unnecessary or improperly billed care (Centers for Medicare and Medicaid Services, 2022).[1] On the other hand, the social costs of excessive oversight are potentially high as well. Distorting healthcare spending could have dire implications for patient health (Doyle et al., 2015). Pressuring providers to cut back spending could deter *necessary* care, especially if it is unclear *ex ante* what services are necessary or not. Given the complexity of identifying healthcare waste, monitoring could also impose considerable compliance costs on providers. If these costs stem mostly from the "back and forth" of the monitoring process, then they pose a deadweight loss that adds to providers' already-high administrative burden (Cutler and Ly, 2011; Dunn et al., 2020). Thus, the extent to which Medicare should monitor depends on the balance between the government savings from increased monitoring and the social costs it imposes on patients and providers.

I study this question in the context of Medicare's largest monitoring program, the Recovery Audit Contractor (RAC) Program. Through the RAC program, private auditing firms ("RACs") conduct manual reviews of individual Medicare claims ("audits") to identify and reclaim payments for unnecessary care. I use novel, audit-level administrative data on RAC audits of hospital stays, Medicare's largest service expenditure category. Four percent of all Medicare hospital admissions were subject to a RAC audit. The rich data in the hospital context on provider operations and patient health offer a unique window into the social costs

[1] Medicare expenditure accounts for 15 percent of total federal spending (Cubanski et al., 2019).

of monitoring. I use hospital cost reports and surveys of hospital technology adoption to characterize hospitals' administrative burden. To examine whether the savings come from reductions in *unnecessary* care, I use emergency department (ED) discharge data that tracks patient outcomes over time, even for patients who are denied a hospital stay.

The empirical analysis brings about three core findings that, taken together, suggest that Medicare is under-monitoring. First, RAC audits have a very high return – every dollar that Medicare spends on monitoring hospitals recovers *$24–29*. The vast majority of these savings stem from the *deterrence* of future spending, rather than the *recovery* of prior spending. Second, monitoring primarily deters *unnecessary* admissions. Hospitals are less likely to admit patients with higher audit risk, and yet these patients were no more likely to return to the hospital due to a missed diagnosis. Third, while provider administrative costs do rise with additional monitoring, the bulk of these costs stem from *investments* that hospitals make. I find evidence of one such investment — hospitals adopt software to assess in real-time whether admitting a patient is medically necessary. This technology adoption in turn leads to persistent reductions in admissions, and as a result the returns to monitoring continue to increase over time. Combining these results with estimates of the value of Medicare spending, I find that increased monitoring is welfare-improving in the medium-run.

The central challenge in identifying the causal effect of monitoring is that RAC audits are endogenous. RACs are private firms that are paid a contingency fee based on the payments they correct. So naturally, they target their audits at claims that are most likely to have an error. I address this endogeneity by leveraging two identification strategies: one compares hospitals subject to differentially aggressive RACs, and the other compares patient cohorts who face exogenously different audit likelihoods.

To understand how hospitals respond to RAC audits, I deploy a difference-in-difference specification comparing hospitals before and after a major expansion of the RAC program in 2011. I focus on hospitals subject to different RACs, leveraging sharp differences in auditing at the border between different RAC jurisdictions. Hospitals subject to a more-aggressive RAC reduce their admissions -– a one percentage point (46 percent) increase in the share of admissions audited leads to a two percent drop in admissions. This effect persists even when auditing is scaled back in later years, consistent with specific deterrence at more heavily-audited hospitals. 89 percent of the savings from the marginal audit stem from the deterrence of future admissions, and the remaining 11 percent are from the payments RACs reclaim. This large deterrence effect is striking, given that policymakers only considered the recovered payments to assess the cost-effectiveness of the RAC program (Centers for Medicare and Medicaid Services, 2012). Extrapolating these effects to the overall hospital sample, I calculate that the RAC program led to upwards of $9 billion in Medicare savings from 2011 to 2015.

Most of the savings from monitoring stem from deterred hospital admissions. In order to identify which patients to no longer admit, I find that hospitals adopted software to aid in the admission decision. Specifically, they are more likely to adopt "medical necessity checking" software, which cross-references electronic health records with medical necessity guidelines set by insurers like Medicare (3M, 2016; Experian Health, 2022). Accordingly, hospital administrative costs rise: for every $1000 in Medicare savings in 2011–2015, hospitals incur $178–218 in administrative costs. But these costs are mostly concentrated as a one-time spike that occurs immediately in 2011. This suggests that provider compliance costs comprise mostly of the fixed costs from investments like technology adoption, rather than the ongoing hassle costs of the monitoring process.

I then turn to investigating the patient health implications of fewer hospital admissions – how well did the reductions target *unnecessary* stays? But because patient composition changes as hospital volume decreases, it is challenging to compare patient outcomes *across hospitals*. In light of this, I exploit a policy which generated exogenous variation in audit likelihoods *across patients* in the same hospital. In particular, I consider the "Two Midnights rule," which barred RACs from auditing patients whose time in the hospital crossed two or more midnights. For this rule, time in the hospital is measured from the point that the patient *arrives* at the ED. Visits that start right after midnight are less likely to reach two midnights than those that start right before. Therefore, patients who arrived at the ED after midnight were more likely to be audited than those who arrived before. I then use a difference-in-difference specification to compare admission rates and health outcomes for before- vs. after-midnight ED patients, pre- and post-Two Midnights rule.

Mirroring the hospital-level results, I find that once the Two Midnights rule is implemented, hospitals cut back on inpatient admissions for after-midnight patients. However, I do *not* find evidence that after-midnight patients were more likely to revisit a hospital within thirty days, a proxy for patient health that is observable in discharge data. Hospitals targeted their reductions to patients in the middle of the severity distribution, who faced up to a 25 percent reduction in admission likelihood. But even among these patients, there is no increase in revisit rates.

Taken together, the empirical results indicate that monitoring providers leads to large returns in terms of Medicare savings, and that the primary tradeoff of these savings comes from provider compliance costs rather than harm to patient health. I then use these estimates to calculate the marginal value of public funds (MVPF) of RAC audits, which reflects society's willingness to pay for each dollar returned to the government by RACs (Slemrod and Yitzhaki, 2001; Kleven and Kreiner, 2006; Finkelstein and Hendren, 2020; Hendren and Sprung-Keyser, 2020). All else equal, a revenue-raising program with a smaller MVPF is "better" as it implies a smaller societal cost of raising a dollar of government revenue.

The MVPF of RAC audits decreases over time, as hospital compliance costs are mostly incurred upfront, but Medicare savings continue to accrue. Three years out, the MVPF is 1.42 and continues to decline after. Assuming a MVPF of Medicare expenditure of 1.63, this implies that RAC auditing is welfare-improving in the medium- to long-run (Finkelstein and McKnight, 2008; Hendren and Sprung-Keyser, 2020). The MVPF calculations also highlight the importance of accounting for the full spectrum of costs and benefits in assessing policy. Failing to account for the savings from deterred care (as policymakers did in their assessment of the RAC program) would suggest that monitoring providers is a very inefficient method of raising revenue. Conversely, ignoring provider compliance costs would make RAC audits seem more cost-effective than they actually were.

This paper contributes to the broad literature on government monitoring and enforcement. Compared to the extensive literatures on monitoring in the context of tax collection, public procurement, and environmental regulation,[2] there has been relatively little work on monitoring in healthcare, despite the large share of public spending that it accounts for. Similar to other contexts, I find that monitoring in healthcare can not only detect, but importantly also deter, unwanted behavior. Interestingly, I find that monitoring can lead healthcare providers to invest in technology to improve cost-effectiveness. This has been a longstanding policy goal that has often proved elusive, despite efforts to directly subsidize such investments (Dranove et al., 2014). Additionally, the relatively high return from using private auditing firms shows that privatized enforcement can be a powerful tool to combat wasteful healthcare spending, mirroring other work looking at using individual whistleblowers to expose fraud (Leder-Luis, 2020; Howard and McCarthy, 2021).

Beyond considering the government savings from monitoring, this paper also makes progress on measuring the *social costs* it imposes. The private costs associated with public programs are often difficult to observe, so their existence is usually deduced indirectly – for example, by looking at how program participation changes when these costs change.[3] I

[2] The baseline theoretical model relating tax enforcement with evasion comes from Allingham and Sandmo (1972), and subsequent extensions to this model and empirical work are surveyed by Andreoni et al. (1998) and Slemrod and Yitzhaki (2002). In the public procurement setting, Olken (2007) shows that top-down government audits of public projects reduces corruption, as measured by the difference between reported and actual expenditure. The empirical literature on environmental regulation has shown that increased monitoring leads to reductions in pollution (Magat and Viscusi, 1990; Hanna and Oliva, 2010; Duflo et al., 2018).

[3] Recent examples include Kopczuk and Pop-Eleches (2007); Deshpande and Li (2019); Finkelstein and Notowidigdo (2019); Meckel (2020); Zwick (2021); Dunn et al. (2021).

leverage a unique context where two forms of social costs – provider administrative costs and patient health outcomes – *can* be observed directly. The MVPF calculations demonstrate how the policy implications are sensitive to the inclusion or exclusion of these costs.

Finally, this paper sheds further light on how healthcare providers respond to incentives. It has been well-documented that providers respond to financial incentives, either by changing the quantity and type of care provided or how they document care.[4] In contrast, less is known about how providers respond to non-financial incentives like monitoring, even though they are employed by both private and public insurers (Gottlieb et al., 2018). This paper contributes to a growing literature on how providers respond to various forms of non-financial incentives: pre-payment denials (Dunn et al., 2021), fraud enforcement (Leder-Luis, 2020; Nicholas et al., 2020; Howard and McCarthy, 2021), and prior authorization (Brot-Goldberg et al., 2021; Roberts et al., 2021).

The rest of the paper proceeds as follows. Section 2 describes the policy context of the RAC program and the data I use. Section 3.1 describes the hospital-level empirical strategy, and Section 3.2 describes the patient-level empirical strategy on ED visits. Section 4 presents the empirical results and incorporates them into a MVPF calculation. Section 5 concludes.

2 Policy Context and Data

2.1 Unnecessary Inpatient Stays and the Recovery Audit Contractor Program

Medicare spent $147 billion, or 19 percent of its total expenditure, on inpatient admissions in 2019 (Medicare Payment Advisory Commission, 2020). Medicare reimburses hospitals a fixed prospective payment per inpatient stay, where the payment depends on the severity-adjusted diagnosis category associated with the stay. Outside of a few exceptions,[5] the payment rate depends on the patient's diagnosis, their pre-existing health conditions, and procedures conducted during their stay. Importantly, it does not generally depend on the admission's length of stay.

Over time, policymakers became increasingly concerned with one area of perceived waste: unnecessary short (0–2 day) stays (Centers for Medicare and Medicaid Services, 2011b; US Department of Health and Human Services Office of Inspector General, 2013). The Medicare

[4] Examples of the former include Cutler (1995); Ellis and McGuire (1996); Clemens and Gottlieb (2014); Einav et al. (2018); Eliason et al. (2018); Alexander and Schnell (2019); Gross et al. (2022); Gupta (2021). Examples of the latter include Silverman and Skinner (2004); Dafny (2005); Sacarny (2018); Gowrisankaran et al. (2019)

[5] One exception is that in "outlier" cases, the payment can depend on length of stay. Outlier stays account for 1.8 percent of overall Medicare hospital stays. Another exception is if an acute care hospital transfers a beneficiary to post-acute care, in which case Medicare pays a per diem rate (Office of the Inspector General, 2019).

Payment Advisory Commission (MedPAC), a non-partisan government agency, contended that hospitals were admitting patients for these short inpatient stays because they were very profitable (Medicare Payment Advisory Commission, 2015): the payment-to-cost ratio for short stays was two times that of longer stays. Appendix Section A.1 describes the Medicare inpatient prospective payment system and short stays in greater detail.

To address this issue, in 2011 Medicare directed RACs to begin monitoring and reclaiming payments for unnecessary inpatient admissions. RAC audits are carried out by four private firms, each of which is in charge of conducting audits within its geographic jurisdiction, or "RAC region." Figure 1a illustrates these regions – they fall along state lines and, in the context of medical claims reviews, are unique to the RAC program.[6] RAC audits were introduced nationally in 2009 after a pilot program in select states. But RAC activity was fairly limited until 2011, when Medicare allowed them to begin auditing unnecessary inpatient stays. The total number of audits increased by *537 percent* from 2010 to 2012, which translated into a *1211 percent* increase in the value of payments reclaimed per hospital (Figure 1b).[7]

Ninety-five percent of inpatient stay RAC audits involve a manual review: the RAC first runs a proprietary algorithm on Medicare claims data to flag individual claims for issues such as missing documentation, incorrect coding, or – starting in 2011 – unnecessary care. A medical professional hired by the RAC, typically a nurse or a medical coder, then requests the documentation for the flagged claim from the provider and manually reviews it. The medical professional determines whether Medicare made an overpayment or, in a small share of cases, an underpayment.[8] If they find an error, then they can demand that the provider repays Medicare (or vice versa). There is no additional penalty to the provider for each corrected payment. The RAC firms are paid a negotiated contingency fee on the payments they correct: 9–12.5 percent, depending on the firm, of the reclaimed payment after appeals. Figure E1 illustrates the full process for claims auditing and appeals, including the remaining 5 percent of inpatient stay audits that do not involve a manual documentation review.

Figure 1b illustrates average per-hospital RAC activity, by year of audit (which is often *after* than the year the claim was originally paid). At the program's peak, RACs were reclaiming $1 million per hospital annually, or 3 percent of the average hospital's Medicare inpatient revenue of $32 million. By 2020, 96 percent of hospitals had at least one inpatient

[6] The RAC regions are also used by Durable Medical Equipment Medicare Administrative Contractors, who do not process claims for medical care, but rather claims for equipment and supplies ordered by healthcare providers. This includes, for example, oxygen equipment, wheelchairs, and blood testing strips.

[7] The total value of reclaimed payments across all hospitals increased from $229 million in 2010 to $3.15 billion in 2012.

[8] In 2011, 6 percent of inpatient stay audits resulted in an underpayment determination.

stay that was audited. RAC audits were then scaled back significantly by 2015, when Medicare paused the program to evaluate complaints made by hospitals and industry stakeholders (Foster and McBride, 2014). Appendix Section A.2 describes the RAC regions, RAC firms, audit process, and timeline of the RAC program in greater detail.

Two years after expanding RAC scope to medical necessity, Medicare introduced a new rule to clarify which admissions could be audited: the "Two Midnights rule." Under this rule, Medicare counted the number of *midnights* during a patient's entire time in the hospital – including the time spent in the ED, in outpatient care, and in inpatient care.[9] If the patient's time in the hospital spanned two midnights, then the stay was presumed to be necessary and RACs could not audit for medical necessity. If the patient's stay *did not* span two midnights, then RACs could audit it (Centers for Medicare and Medicaid Services, 2017). So for the 73% of Medicare inpatient admissions that originate in the ED, the Two Midnights rule effectively increased audit likelihoods for patients who arrived after midnight, relative to those who arrived before.

2.2 Data

The hospital-level analysis uses four main data sets. First, I use audit-level administrative data on the RAC program acquired through a Freedom of Information Act request. The data span 2010 to 2020 and include claim-specific information on 100 percent of RAC audits, such as characteristics of the audited claim (e.g., hospital, admission date, discharge date, diagnosis, Medicare payment) and of the audit (e.g., audit date, audit decision, amount of payment reclaimed or corrected, appeals). The dataset covers 4.5 million audits of inpatient stays.

Second, I use Medicare inpatient and outpatient claims data. I merge the RAC audit data with the Medicare inpatient claims data (Medicare Provider Analysis and Review; MEDPAR) by matching on the following elements: provider, admission and discharge dates, diagnosis-related group, and initial payment amount. I am able to identify whether a claim was audited for 99.6 percent of Medicare inpatient claims between 2007 and 2015. I also conduct analyses using Medicare outpatient claims to measure the use of observation stays and total outpatient revenue.

[9]Midnight cutoffs are surprisingly common in insurer billing rules; see the policies studied by Almond and Doyle (2011) and Rose (2020). A difference between the Two Midnights rule and the policies studied by Almond and Doyle (2011) and Rose (2020) is that the Two Midnights rule counts the number of midnights during a patient's entire stay in the hospital, starting from when they arrive at the hospital. In contrast, the rules studied by these two papers focus on how many midnights pass during a patient's hospital admission, starting from the *hospital admission hour* (that is, the hour that the patient is formally admitted for inpatient care or, in the case of newborns, born).

Third, I use hospital cost data from the Healthcare Cost Report Information System (HCRIS), which collects cost reports that hospitals submit to Medicare. In particular, HCRIS provides yearly measures of hospital administrative costs.

Fourth, I use data on IT adoption from the Healthcare Information and Management Systems Society (HIMSS) Analytics Database, which is a yearly survey of IT used by hospitals and other healthcare providers. HIMSS asks hospitals each year to report the types of IT they are planning to or have already installed. In particular, I focus on medical necessity checking software, which hospitals use to assess the medical necessity of care in real-time. Additionally, to study heterogeneity across hospital types, I also use hospital characteristics from the Medicare Provider of Services file and hospital group affiliations from Cooper et al. (2019).

Table 1 presents summary statistics by RAC region. Hospitals in Regions B (Midwest) and C (South) have much lower audit rates than hospitals in Regions A (Northeast) and D (West). Within each region, rural hospitals, small hospitals, non-profit hospitals, and hospitals with a higher share of short stay Medicare admissions are more likely to be audited (Figure E4). Appendix Section A.3 further explores the claim-level and hospital-level characteristics associated with auditing in further detail.

In the patient-level analysis of ED visits, I use the Florida State Emergency Department Database (SEDD) and State Inpatient Database (SID) between 2010 and 2015. I focus on Florida because it is the only state that reports ED arrival hour in the publicly available data for both the inpatient *and* emergency department datasets; Medicare's Inpatient and Outpatient files do not report this variable.[10] The most granular unit of time for ED arrival in my data is the hour. SEDD includes discharge-level data on every outpatient ED visit, and SID includes every inpatient stay (and denotes whether the patient was admitted as inpatient from the ED). I combine the two to construct the universe of ED visits in Florida hospitals in this time period. I proxy for patient health after an ED visit by considering whether the patient revisits any hospital in Florida shortly after, either as an ED visit or an inpatient visit.[11] I use this proxy because mortality is not observable in hospital discharge data such as SID and SEDD. Table F3 presents patient characteristics common

[10]ED visits are known to be difficult to identify using claims data, as there is no standard method or definition. For example, whether a patient who receives an ED triage evaluation without emergency clinician professional services (e.g., evaluation by a primary care clinician) is considered an "ED visit" has been found to vary across different data sources (Venkatesh et al., 2017). Further, in my attempt to assemble a panel of ED visits using Medicare claims, I uncovered inconsistencies in the data that, after consulting with ResDAC, lead me to conclude that across-year and across-provider comparisons of ED visits are untenable using the Medicare claims (ResDAC, 2022).

[11]Hospital inpatient readmission rates are a widely used measure of hospital quality (Krumholz et al., 2017). Reducing hospital readmissions was the focus of the Hospital Readmissions Reduction Program, one of the value-based purchasing programs introduced as part of the Affordable Care Act.

across MEDPAR and SID/SEDD, and compares the overall inpatient sample (MEDPAR), border hospital inpatient sample (MEDPAR), inpatients admitted from a Florida ED (SID), and patients admitted from a Florida ED who arrived at the ED within 3 hours of midnight (SID). The samples are similar in terms of age, sex, race, and share with a recent inpatient stay.

Table 2 reports summary statistics for before- and after-midnight arrivals before the Two Midnights rule (in 2013Q2). Figure 2 plots the quarterly share of before- and after-midnight Medicare ED arrivals who are admitted as inpatient. Prior to the Two Midnights rule, after-midnight arrivals are more likely to be admitted as inpatient, but this gap closes once the Two Midnights rule is implemented in 2013Q3.

3 Identification Strategies

3.1 Identifying the Effect of Monitoring on Hospital Outcomes

The aim of the first, hospital-level identification strategy is to understand how hospital behavior responds to the increase in auditing in 2011. I focus on hospitals close to the RAC border and compare hospitals who are subject to a more-aggressive RAC to their neighbors who are subject to a less-aggressive one. I then look at how their behavior changes after 2011 using a difference-in-difference specification, with two modifications. First, I include local fixed effects to compare hospitals that are neighbors to each other. Second, I instrument for a hospital's audit rate using a measure of how aggressively its RAC audits *other* hospitals.

Border Hospital Sample: Figure 1a illustrates the sharp changes in audit intensity at the border between RAC regions. The changes across the RAC borders are twice as large as the changes across state borders *within* each RAC region. I focus on the sample of hospitals close to the border, where I define "close" as being within one hundred miles of it. By focusing on this subset of hospitals, this research design requires a weaker parallel trends assumption relative to one incorporating all hospitals. Here, I only need to assume that *geographically proximate* hospitals are not on differential trends.

Neighbor Comparison Groups: To ensure that I am comparing hospitals that are close *to each other*, and not just hospitals that are close to the border, I identify a unique set of neighbors for each hospital and call this its "neighbor comparison group."[12] I define a hospital's neighbor comparison group to be hospitals on the *other* side of the border within 100 miles. I include a fixed effect for each group interacted with a year indicator in my

[12] In identifying a unique set of neighbors for each hospital, I follow Dube et al. (2010), whose state border-county identification strategy allows individual counties to be paired with unique sets of adjacent counties.

specification. With these fixed effects, I effectively "stack together" local comparisons of hospitals to their neighbors across the border.

Figure E5 illustrates how I construct a hospital's neighbor comparison group. The hospital in question is on the Oklahoma side of the border (RAC Region C) and has an audit rate of 1.44 percent. The members in its neighbor comparison group are the hospitals on the other side of the border within a hundred miles – in this case, that would be hospitals in Kansas (RAC Region D) that face a much higher average audit rate of 5.42 percent. Together, the Oklahoma hospital and its neighbors in Kansas form the neighbor comparison group for the Oklahoma hospital.

Including these group-year fixed effects improves upon a specification with just border fixed effects (or border-year fixed effects) in two ways. First, it accounts for local geographic trends in utilization and spending. Prior work in the healthcare literature has documented substantial geographic variation in Medicare spending (Skinner, 2011; Finkelstein et al., 2016). Each RAC border spans hundreds of miles. A specification with just border fixed effects would therefore end up comparing hospitals that are close to the border, but possibly far from *each other*; this may not adequately account for local trends. Second, constructing these neighbor comparison groups allows me to include hospitals at the intersection of multiple borders. In a specification with border fixed effects, I would have to either arbitrarily assign these hospitals to one of their adjacent borders, or exclude them from the analysis.

Because a hospital can be a member of multiple neighbor comparison groups, the sample includes repeated hospital observations which will have correlated errors. To account for this, I divide the border into segments and cluster at the border segment level. Figure E6 illustrates the border segments used for clustering, with each segment in a different color. Each border segment is a hundred miles, except for segments that cross state lines, which are split at the state border.

Event Study Specification: The event study specification of interest for the hospital-level strategy is:

$$Y_{ht} = \sum_{\tau=2007}^{2015} \mathbb{1}[t = \tau] \times X_h^{2011} \beta^\tau + \phi_{g(h)t} + \psi_h + \varepsilon_{ht} \ . \tag{1}$$

In Equation 1, Y_{ht} is an outcome for hospital h in year t, X_h^{2011} is the hospital's 2011 audit rate, $\phi_{g(h)t}$ is a hospital's neighbor comparison group $g(h)$-times-year fixed effect, and ψ_h is a hospital fixed effect. The main results are in the form of an event study to allow for dynamic responses, so there is a β^τ for each year τ between 2007 and 2015, omitting 2010. β^τ can be interpreted as the effect of a one percentage point increase in 2011 audit rate on a hospital outcome in year τ, relative to 2010.

Audit Rate Instrument: One concern with estimating Equation 1 directly is the endogeneity of a hospital's 2011 audit rate X_h^{2011} – that is, that $E[\varepsilon_{ht}|X_h^{2011}] \neq 0$. This could arise if hospitals that are targeted by RACs were on a differential trend relative to their neighbors – for example, if RACs target lower-quality hospitals and admissions at lower-quality hospitals were already on a downward trend. To isolate variation driven by the *RAC* and not by the hospital, I consider how aggressively the RAC audits *other hospitals* under its jurisdiction. In practice, I instrument for a hospital's 2011 audit rate with the audit rate of other hospitals in the same state. For each hospital, I calculate the "leave-one-out state audit rate," which is the state average *excluding* that hospital. It is defined as:

$$Z_h^{2011} = \frac{1}{n_{s(h)} - 1} \sum_{h' \in \ s(h) \setminus h} X_{h'}^{2011} \ , \tag{2}$$

where $X_{h'}^{2011}$ is the 2011 audit rate for hospital h' that is in the same state $s(h)$ as hospital h. Because RAC borders fall along state lines, hospital h' is subject to the same RAC as hospital h. There are $n_{s(h)}$ total hospitals in the state.

The reduced form event study specification is:

$$Y_{ht} = \sum_{\tau=2007}^{2015} \mathbb{1}[t = \tau] \times Z_h^{2011} \gamma^\tau + \phi_{g(h)t} + \psi_h + \varepsilon_{ht} \ . \tag{3}$$

In order to interpret the coefficients as the effect of a one percentage point increase in the 2011 audit rate (as in Equation 1), I scale the γ^τ coefficients in Equation 3 by the correlation between X_h^{2011} and Z_h^{2011} (after accounting for hospital-group fixed effects).[13]

I also report results that pool the post-2011 effects into a single coefficient:

$$Y_{ht} = \mathbb{1}[t \geq 2011] \times X_h^{2011} \beta^{post} + \phi_{g(h)t} + \psi_h + \varepsilon_{ht} \ . \tag{4}$$

In this case, the reduced form specification is:

$$Y_{ht} = \mathbb{1}[t \geq 2011] \times Z_h^{2011} \beta^{post} + \phi_{g(h)t} + \psi_h + \varepsilon_{ht} \ . \tag{5}$$

Identification Assumptions and Checks: The identification strategy relies on three underlying premises: first, that the changes in audit rate at the border are driven by RACs

[13] In particular, I generate eight instruments, each of which is an interaction of Z_h^{2011} with a year indicator, and combine them to instrument for the interactions of X_h^{2011} with a year indicator. For example, I use $\sum_{\tau=2007}^{2015} \mathbb{1}[t = \tau] \times Z_h^{2011}$ to instrument for $\mathbb{1}[t = 2007] \times X_h^{2011}$, and the coefficient is equal to the correlation between X_h^{2011} and Z_h^{2011} when $\tau = 2007$,and zero for $\tau \neq 2007$. I repeat this for all 8 years between 2007 and 2015. I implement this in a two-stage procedure to allow for clustering in the estimation of standard errors.

(*exogeneity*); second, that neighboring hospitals are "comparable" to each other (*parallel trends* and *homogeneous treatment effect*); and third, that the leave-one-out state audit rate is a valid instrument for the hospital audit rate (*exclusion restriction* and *monotonicity*).

First, suppose that the sharp changes in audit rate at the border in Figure 1a were *not* driven by variation across RACs. If they were instead driven by hospital or patient characteristics (or a policy that is correlated with them) we would expect to see similarly sharp variation at the border in these characteristics as well. Figure E8a plots a hospital-level measure that is highly correlated with 2011 audit rates in the cross-section: the short stay share of 2010 Medicare admissions. Figure E8b plots the predicted 2011 audit rate, where the prediction depends on patient stay characteristics but *not* the identity of the RAC. Neither of these measures displays sharp changes at the border, suggesting that the pattern in Figure 1a is not driven by the hospitals or patients themselves, but instead by the RACs and the RAC regions.

On each side of the border, RACs face the same incentives to audit and presumably similar local labor costs. So what could be driving these sharp differences in audit rate across the RAC border? One explanation could be that because each RAC comes from a different industry background,[14] this variation in prior experience translates into differences in how RACs approach auditing. These differences would be especially pronounced in 2011, as it is the first year that RACs were allowed to conduct medical necessity audits. Another explanation could be that RACs set their audit strategies at the regional, rather than local, level. For example, this would be the case if RACs combined data from all hospitals in its region to train a single algorithm to flag claims, so a hospital's audit rate would reflect within-region spillovers via the algorithm. Or, it could be that RACs set their audit rates based on the average *regional* labor cost of hiring auditors, rather than the local labor cost.

Second, the border hospitals must be "comparable" to each other. Note that I do not need to assume there are *no differences* in hospitals across the RAC border – this would be clearly violated by the fact that hospitals on opposite sides of the border are in different states. Instead, I need to make weaker assumptions: that hospitals on each side of the border have parallel trends and homogeneous treatment effects. With the inclusion of group-year fixed effects, for the parallel trends assumption we only need that neighboring hospitals on opposite sides of the border do not differentially deviate from local trends. While this assumption is in principle untestable, a lack of preexisting differential trends in the event study would support it.[15] The parallel trends assumption could be violated if the results

[14]For example, the RAC in Region A is primarily a debt collection agency, while the RAC in Region C is a healthcare data analysis company.

[15]Restricting the comparison to *border* hospitals allows me to make a weaker parallel trends assumption than a comparison of *all* hospitals. Figure E15f shows the results from an alternate specification that includes

are actually driven by state policies changing over time. In robustness tests I show that the results are robust to omitting individual states, and therefore are not driven by any individual state's policy changes. Thus in order for the results to be driven by state policies, the policy changes would have to be consistent across *multiple* states on one side of the border, and they would all have to change simultaneously in 2011.

Since a hospital's audit rate is continuous and therefore "fuzzy," I also need to assume that hospitals in the border sample have homogeneous treatment effects (de Chaisemartin and D'HaultfŒuille, 2018). One concern is that if hospitals on opposite sides of the border are very different at baseline, then they may also have heterogeneous responses to auditing. Table F2 reports the correlation between 2010 hospital characteristics and audit rates in the border hospital sample and the overall sample. Comparing *within* neighbor comparison groups for the border hospital sample, the 2011 audit rate is uncorrelated or weakly correlated with 2010 hospital characteristics. In contrast, these correlations are statistically significant and larger in magnitude in the overall sample.

Finally, to justify using the leave-one-out state audit rate as an instrument, I need the exclusion restriction as well as a monotonicity assumption. The exclusion restriction requires that the leave-one-out audit rate only affects a hospital's outcomes via its own audit rate. To violate this, time-varying confounders like changes in state policies would have to be consistent across multiple states and occur simultaneously in 2011. Non-time-varying confounders like existing state policies are absorbed by the hospital fixed effect in the difference-in-difference specification. The exclusion restriction could also be violated by reverse causality – if, say, the leave-one-out audit rate reflects a given hospital's spillovers onto other hospitals in the same state. This could be true if a given hospital has a large market share, or if hospitals in the same chain have spillovers on each other. To address this concern, I run robustness tests that instrument using the average audit rate of hospitals in the same state but in other markets, as well as hospitals in the same state but not in the same chain. The results from using each of these instruments are similar to the main results (Figure E16). Additionally, note that we need to make an assumption about monotonicity in audit intensity across RACs – that a given hospital would be subject to more audits under a more-aggressive RAC, and fewer audits under a less-aggressive RAC (Imbens and Angrist, 1994).

3.2 Identifying the Effect of Monitoring on Patient Outcomes

I next turn to the patient-level identification strategy that leverages the Two Midnights rule. I split ED visits by whether the patient arrived before midnight (lower audit risk)

all hospitals; there is evidence of differential pretrends when comparing across all hospitals.

or after midnight (higher audit risk), and then compare them pre– and post–policy in a difference-in-difference specification.

Specification: The event study specification is:

$$Y_v = \sum_{\tau=2010Q1}^{2016Q4} \mathbb{1}[q=\tau] \times \mathbb{1}[t \geq 00{:}00]\beta^{\tau} + \boldsymbol{W_v'\gamma} + \lambda_{hq} + \phi_{ht} + \varepsilon_v \ , \tag{6}$$

where ED visit v occurs in quarter q at hospital h, and the ED arrival hour of the visit is $t \in [21{:}00, 03{:}00)$ (that is, between 9PM and 3AM). Y_v is the outcome of interest, such as an indicator for whether the ED visit resulted in an inpatient admission or whether the patient revisited a hospital within thirty days. $\mathbb{1}[q=\tau]$ is an indicator for whether the visit occurred in quarter τ, omitting 2013Q3. $\mathbb{1}[t \geq 00{:}00]$ is an indicator for whether the patient arrived at the ED after midnight. λ_{hq} is a hospital-quarter fixed effect, and ϕ_{ht} is a hospital–ED-arrival-hour fixed effect. W_v are controls for patient characteristics, including patient age, race, Hispanic, point of origin, an indicator for whether last ED visit was within 30 days, number of chronic conditions, and average income in patient's zip code. β^{τ} is the coefficient of interest and can be interpreted as the effect of the increased audit likelihood on after-midnight ED arrivals in quarter τ, relative to 2013Q3.

Equation 7 pools the event study into a single post-policy coefficient β:

$$Y_v = \mathbb{1}[q \geq 2013Q3] \times \mathbb{1}[t \geq 00{:}00]\beta + \boldsymbol{W_v'\gamma} + \lambda_{hq} + \phi_{ht} + \varepsilon_v \ . \tag{7}$$

Here $\mathbb{1}[q \geq 2013Q3]$ is an indicator for whether the visit occurs after the Two Midnights rule is implemented in 2013Q3.

Identifying Assumption and Checks Interpreting β and β^{τ} as the causal effects of auditing requires two assumptions. First is the standard parallel trends assumption – that absent the Two Midnights rule, before- and after-midnight patients would have trended similarly. To substantiate this, I check that there are no differential pre-trends between the two groups in the event study figures.

The second assumption is that there is no manipulation of the ED arrival hour. This would be violated if, for example, hospitals misreported after-midnight ED arrivals as arriving before midnight. If this were the case, we would expect to see bunching of ED arrivals right before midnight once the policy is implemented (that is, an increase in the share of patients reported arriving between 11:00 PM and midnight). Figure E9 plots the share of patients by ED arrival hour, pre- and post-policy – bunching before midnight does not appear post-policy. I test this empirically in Table F4 by looking at whether there is a higher share of patients arriving in the hour before midnight (column 1) or a lower share of patients arriving after midnight (column 2) post-policy. Neither of these measures changes after the Two Midnights rule is implemented.

Practically speaking, it may be difficult for hospitals to manipulate the ED arrival hour in response to the Two Midnights rule. The arrival hour is recorded as soon as the patient walks in to the ED, which makes it more difficult to manipulate than a measure that is recorded later on. Additionally, to game the Two Midnights rule, hospitals would have to make after-midnight arrivals look like before-midnight ones. This would require them to actively *move up* a patient's ED arrival hour to an earlier time, rather than a more passive form of misreporting by "dragging their feet" to record a later arrival hour, in contrast to other contexts where this kind of behavior has been found (e.g., Chan (2016); Jin et al. (2018)).

We may also be concerned that hospitals respond to the Two Midnights rule by simply extending all stays to span two midnights. This would not be a threat to identification per se; instead we would simply see no effect of the Two Midnights rule on inpatient admission likelihood. Due to patient confidentiality reasons in the discharge data, I cannot directly observe how long a patient's entire stay in the hospital spanned. However, I do not find evidence that after-midnight patients have additional charges, diagnoses, or procedures after the rule is implemented (Table F5), suggesting that hospitals did not respond to the Two Midnights rule by extending stay duration.

4 Results

4.1 Hospital Outcomes: Admissions, Revenue, Costs, and IT Adoption

Figure 3 plots a binscatter of the cross-sectional relationship between the instrument, the leave-one-out state audit rate, and hospital audit rates in the border hospital sample. The leave-one-out audit rate explains 74 percent of the variation in the actual audit rate, with a coefficient of 1.04. There is a positive linear relationship between the two and it is not driven by outliers, which supports using a linear specification.

Figure 4 presents the first set of main results: the event study coefficients on hospital-level outcomes, scaled by the cross-sectional correlation between the audit rate and the leave-one-out audit rate in Figure 3. Table 3 reports the yearly coefficients for 2011 to 2015 (for brevity, the pre-2011 coefficients are estimated but not reported in the table). Figures 4a and 4b plot the results for log Medicare admissions and log Medicare inpatient revenue, where inpatient revenue is defined as the sum of all Medicare inpatient payments. Prior to 2011, hospitals with higher audit rates do not seem to be on differential trends relative to their neighbors across the border. Starting in 2011, there is a decline and then a plateau in Medicare admissions and inpatient revenue among hospitals subject to a more-aggressive RAC. A one percentage point increase in the 2011 audit rate results in a 1.1 percent decrease in admissions in 2011, which increases in magnitude to a 1.9 percent decrease by 2012 and 2013. Similarly, a one percentage point increase in the 2011 audit rate results in a 1.0 percent decrease in inpatient revenue in 2011, and then a 1.8 percent decrease by 2012 and a 2.8 percent decrease by 2013. Extrapolating these estimates to the overall hospital sample (albeit under fairly strong assumptions) indicates that RAC audits saved the Medicare program \$9.28 billion between 2011 and 2015.[16]

I next turn to the administrative burden RAC audits place on hospitals. Figure 4 and Table 3 columns 5-6 present results on two dimensions of this burden: hospital administrative costs and IT adoption. Figure 4c plots estimates of the effect on log administrative costs, as reported in hospital cost reports. A one percentage point increase in RAC auditing in 2011 results in an immediate 1.5 percent uptick in administrative costs, but this increase lasts for only about a year. This result corroborates the findings of an AHA survey in which 76 percent of surveyed hospitals reported that RAC audits increased their administrative burden (American Hospital Association, 2012).

Investments into technology to improve compliance and mitigate monitoring can be a driver of higher administrative costs. A particularly relevant type of technology is "medical necessity checking software," which hospitals use to assess the medical necessity of the care they provide with respect to payer coverage rules (3M, 2016; Experian Health, 2022). Figure 4d presents the event study results for whether a hospital reported installing medical necessity checking software in a given year. In response to a one percentage point increase in the 2011 audit rate, hospitals were 2.2 percentage points more likely to report that they were installing or upgrading this software in 2012 (relative to the 59 percent of hospitals that had this software installed in 2010).

Given policymakers' concerns about short stays being the primary driver of unnecessary stays , Figure 5 splits admissions by their length of stay. The overall reduction in admissions is driven by a reduction in short stays – that is, admissions with length of stay less than or equal to two days. A one percentage point increase in the audit rate results in a 4.6 percent decrease in short stay admissions and a 4.6 percent decrease in revenue from these stays by 2012 (Table 3). In contrast, there is a much smaller and statistically insignificant decrease in longer stay admissions. Figure E10 splits admissions by their Major Diagnostic Category (MDC) to investigate which diagnoses are driving the reduction in admissions. In particular, I split admissions for whether or not they are in the circulatory system MDC, motivated by

[16]The procedure and assumptions underlying this extrapolation are discussed in Appendix Section E.

the large number of audits for circulatory system diagnoses (Figure E2).[17] A one percentage point increase in the audit rate results in a 4.8 percent decrease in circulatory diagnosis admissions and a 5.9 percent decrease in revenue from these stays in 2012; in contrast, it only leads to a 1.2 percent decrease in non-circulatory diagnoses.

Figure E11 plots the results for the payments directly reclaimed by RACs. A one percentage point increase in audit rate in 2011 is associated with $314,115 in demanded payments in 2011 per hospital. There are additional demands in subsequent years as well, although the magnitude diminishes over time. Comparing the savings from deterred admissions to reclaimed payments, I calculate that 89 percent of government savings from the RAC program are due to deterrence. Overall, RAC auditing brings in $24 in Medicare savings per dollar spent to run the program.[18]

Figure E12 considers whether hospitals substituted away from inpatient care to outpatient care – for example, to observation stays. Observation stays consist of short-term (often diagnostic) services provided at the hospital while a physician decides whether to admit a patient or send them home.[19] At the hospital-level, I find no evidence that hospitals increased outpatient spending or observation stays in response to RAC audits.

The event studies in Figure 4 also illustrate the dynamics of hospital responses. Admissions and revenue decline steadily between 2011 and 2012. The fact that this happened over two years rather than immediately likely reflects two factors. First, some of the 2011 admissions occurred before hospitals knew how aggressively they would be audited by RACs. Second, it may have taken time to implement practices or technology to reduce unnecessary admissions. After 2012, admissions remained at their decreased levels – even in 2014 and 2015, when audit activity slowed down significantly. In contrast, there was an immediate but short-lived increase in hospital administrative costs in 2011. The timing of this effect suggests that the bulk of hospital compliance costs were *fixed*, rather than *variable*, costs. If

[17] For example, the "Chest Pain" DRG (313) has the highest improper payment rate (30.5%) among all DRGs (Centers for Medicare and Medicaid Services, 2018).

[18] For a one percentage point increase in 2011 audit rate, the government costs by 2015 are $88k, savings from reclaimed payments are $232k, and the total Medicare savings are $2.08 million. These numbers are calculated under the assumption that CMS *returned* 68 percent of reclaimed payments to hospitals. I assume this because in August 2014, Medicare announced a one-time option to return part of the reclaimed payments in exchange for hospitals dropping their appeals. See Section for more details on the settlement. Under the assumption that hospitals do *not* settle and Medicare keeps all the payments they demand, the savings by 2015 from reclaimed payments are $721k, and total government savings are $2.57 million. Thus in this case, RAC audits save $29 per dollar of monitoring costs, and deterred admissions account for 72 percent of the savings.

[19] Observation stays typically last less than forty-eight hours and are billed as an outpatient service. They are often cited as a more cost-effective alternative to a short inpatient stay (Medicare Payment Advisory Commission, 2015). Since observation stays occur in the hospital and many hospitals do not have separate observation units, patients often are not aware they are in an observation stay rather than a formal inpatient stay (Span, 2012).

the costs were primarily variable costs, then we would expect to see elevated costs for several years, since audits continued until 2015 (Figure 1b). Instead, the one-time spike in administrative costs is consistent with hospitals making upfront investments to improve compliance going forward; the installation of medical necessity checking software is an example of one such investment.

The results also suggest that prior to 2011, the high rate of short stays was not necessarily due to hospitals *knowingly* admitting unnecessary admissions. If this was the case, hospitals would not have needed to install technology in order to assess the medical necessity of care. One might also expect them to only cut back on admissions while RACs are active, and quickly ramp them up again when they are less active. On the contrary, I show that being exposed to a high audit rate in 2011 has a persistent deterrence effect, even in later years when audit levels are much lower.

Table F6 pools the post-2011 years of the main results into a single post-2011 coefficient, as in Equation 5. Given the dynamics of the results, the pooled coefficients are noisily estimated. Averaging across 2011 to 2015, there is a 1.5 percent reduction in overall admissions (although not statistically significant) and a 2.2 percent reduction in short stay admissions relative to the pre-period. Table F7 considers heterogeneity in the effect by hospital characteristics. Rural, for-profit, smaller, and non-chain hospitals are more responsive to audits. Reassuringly, the increase in medical necessity checking software is driven by hospitals that do not have the software installed in 2010. Appendix Section B checks that the results are robust to instrumenting for the share of claims that are *denied* rather than just audited, using varying bandwidths to define the hospital sample, excluding hospitals that are very close to the border, using alternative instruments for audit rate, removing individual states or neighbor comparison groups, and running a placebo test using state borders in the interior of each RAC region. In Appendix C, I consider whether RAC audits affected rural hospital closure rates in subsequent years. If hospitals lost enough revenue from auditing that it caused them to close, then this would have important implications for patient access to care. Figure E13 shows that border hospitals subject to more auditing were no more likely to close in subsequent years, mitigating concerns about this channel.

Overall, the hospital-level analysis shows that auditing saved money for Medicare primarily by deterring unnecessary admissions, but the burden of identifying which admissions to cut back on fell on hospitals. A back-of-the-envelope calculation comparing the total government savings with the compliance costs finds that for every \$1,000 in savings between 2011 and 2015, hospitals spent \$218 in compliance costs.[20]

[20] The value of compliance costs by 2015 is \$455k, compared to the total government savings of \$2.08 million. Under the assumption that a hospitals do not settle and CMS does not return reclaimed payments

4.2 Patient Outcomes: Inpatient Admission Likelihood and Revisit Likelihood

Figure 6 plots the event studies of the patient-level analysis of ED visits in Equation 6. There is no clear trend in the pre-policy coefficients, which supports making the parallel trends assumption. Immediately after the Two Midnights rule is implemented, there is a drop in the share of after-midnight ED arrivals that result in an inpatient admission. There is a symmetric increase in the share of patients who are not admitted, but are placed into observation.

Table 4 reports the β coefficient from Equation 7. In columns 1 and 2, the coefficients on the inpatient indicator and observation indicator are symmetric in opposite directions. After the Two Midnights rule goes into effect, after-midnight arrivals are 0.7 percentage points (1.7 percent) less likely to be admitted as inpatient and 0.7 percentage points (14 percent) more likely to be placed in observation. There is no change in the share of patients who are sent home directly from the ED ("Not Admitted"). This indicates that for ED patients who are on the margin for being admitted as an inpatient, hospitals still preferred to keep them in the hospital rather than sending them home directly.

Next, I consider whether the reduction in inpatient admissions harmed patients. Panel 6d plots the event study results for an indicator of whether a patient revisited a hospital within thirty days of her ED visit, and column 4 in Table 4 reports the pooled coefficient. Despite their reduced inpatient admission rate, there was no increase in revisits for after-midnight patients. This finding is in line with other work which has found that the marginal hospitalization has no effect on mortality (Currie and Slusky, 2020). However, because only a small subset of patients should be on the margin of an admission, this null average effect may be masking heterogeneity across patients. Patients in the middle of the severity distribution should be more likely to be denied admission as a result of RAC audits, so one would also expect any health effects to be concentrated among these patients as well.

To explore this heterogeneity, I predict a patient's severity based on information available at the outset of an ED visit. Using data on ED visits between 9:00 AM and 3:00 PM (that is, a time window outside of that used for the main results), I estimate a logistic regression predicting whether a patient is admitted within thirty days of the visit, based on information available during an ED visit.[21] I then apply this prediction to the main sample to create a measure of predicted patient severity, and split patients into deciles of this measure. I

to hospitals, the total government savings are \$2.57 million, so the ratio between compliance costs and savings is \$177 in hospital compliance costs per \$1000 in Medicare savings.

[21]This includes patient demographics such as age-bin, sex, race, a Hispanic indicator, a point-of-origin indicator, and mean zip code income. It also includes hospital and quarter fixed effects; the number of visits, inpatient stays, or length of stay in the last month or last year; and any diagnoses and procedures recorded for stays within the last month or last year.

reestimate the specification in Equation 7, interacting β with an indicator for each severity decile.

Figure 7 plots the heterogeneity by severity results for inpatient status and for revisits within thirty days, and Table F10 reports the coefficients. The Two Midnights rule has no effect on admission rates for patients at the bottom and top severity deciles. Instead, the reduction in admissions is coming primarily from the middle of the severity distribution. There is a 5 percentage point, or 25 percent, decrease in admissions for patients in the fifth predicted decile. However, I do not see this pattern for revisits, as the coefficient on revisits is statistically insignificant at all risk deciles. Thus, the overall null effect on revisits is not masking heterogeneity by patient severity. Even among patients with the highest likelihood of being denied admission, there is no increase in revisits.

Table F8 reports heterogeneity of the patient-level effect by hospital characteristics. Urban, teaching, for-profit, and smaller hospitals are more responsive to the rule. Notably, the response is driven by hospitals with the medical necessity checking software in 2012. This speaks to the usefulness of this software – it could help hospital decided whether to bill as an inpatient or observation stay by notifying them of billing rules such as the Two Midnights rule. Appendix Section B shows that the results are robust to varying the bandwidth used to define before- and after-midnight ED arrivals, the period used to measure hospital revisits, as well as a falsification test on non-Medicare patients, who should not be directly affected by the Two Midnights rule.

Both the hospital-level and patient-level approaches find a decrease in admissions, but one difference between the two sets of results is that only the patient-level approach finds a symmetric increase in *observation* stays. This could be due to differences in the measurement of observation stays across different data sources. For the patient-level analysis on ED visits, I follow the Agency for Health Care Administration (2015) and define an observation stay as a visit with a charge for observation services that is not part of an inpatient admission. I use a similar definition in the Medicare outpatient claims (albeit not restricting to ED visits),[22] but recent work by Sheehy et al. (2019) and Powell et al. (2020) finds that this may overcount observation stays in the Medicare data.[23] The different observation stay results could also reflect differences in the patient sample captured in each approach: the patient-level analysis subsets to ED visits, whereas the hospital-level analysis considers all admissions. Part of the

[22] I count any outpatient claim with an observation revenue center ("0760" or "0762") or procedure code ("G0378" or "G0379") as an observation stay.

[23] Sheehy et al. (2019) finds evidence that some "observation stays" in the Medicare outpatient claims should actually be counted as part of an inpatient stay, and tests various methods to identify when this is the case. Powell et al. (2020) find that a significant portion of these "observation stays" seem to be for preplanned, repeated services to treat chronic conditions.

decrease in admissions at the hospital level could reflect efforts to deter patients *before* they ever arrive at the hospital – for example, by discouraging physician referrals and transfers, influencing ambulance referral patterns, or deciding to not expand ED capacity.

4.3 Marginal Value of Public Funds Calculation

The results suggest that most of the impact of RAC audits comes from its effect on government savings and hospital compliance costs, rather than any changes in the quality of care patients receive. We can use the estimates to assess the effect of a *marginal* increase in audit rate: the marginal value of public funds (MVPF) of the RAC program. The MVPF is the ratio of society's (i.e., hospital and patient) willingness to pay for each additional dollar that gets returned to the government budget through the RAC program.

Marginal value of public funds:

$$\text{MVPF} = \frac{\Delta \text{ hosp. revenue} + \Delta \text{ hosp. compliance costs} + \Delta \text{ treatment costs} + \Delta \text{ pt health}}{-\Delta \text{ hospital revenue} + \Delta \text{ monitoring costs}} . \quad (8)$$

The numerator of the MVPF is the societal willingness to pay to *avoid* an increase in auditing. This depends on the change in hospital revenue, the change in hospital compliance costs, the change in treatment costs, and the change in patient health. The denominator of the MVPF is the change in government budget due to an increase in auditing. This depends on the change in hospital revenue (which is equal to Medicare's savings) and the change in government monitoring costs.

Given the the dynamics of hospital responses, the time horizon considered is important. If hospitals incur fixed costs such as a large upfront investment in technology, then these costs should be compared to the discounted value of savings over a multiyear horizon. To remain agnostic about the time horizon for calculating welfare, I calculate the MVPF of RAC audits using the cumulative costs and benefits for each year between 2011 and 2018.

The assumptions and parameters used for the baseline MVPF calculation are listed in Table 5, and the details of the calculation are discussed in Appendix Section D. I use the estimates derived from the event study in Figure 4 and Table 3 to inform the effect on hospital revenue and hospital compliance costs. To calculate the effect on government monitoring costs, I multiply the reclaimed payments in Figure E11 by RACs' contingency fees. At baseline, I assume a contingency fee of 10.75 percent (the average of the range of contingency fees, from 9 and 12.5 percent). For the value of the change in patient health, I assume in the baseline calculation that it is 0. This is motivated by the null result from the patient-level results.[24]

For the change in treatment costs, I assume at baseline that this does not change and is equal to 0. This would be the case if hospitals substituted inpatient admissions with other care that has the same cost, or used the freed up capacity to treat non-Medicare patients more intensively. In practice, this assumption is likely a lower bound on the treatment cost savings. If hospitals instead incurred *lower* treatment costs after reducing admissions, then the societal savings would be even larger and thus the MVPF would be smaller.

Figure 8a plots the yearly MVPF of an increase in the 2011 audit rate. The MVPF in 2011 is relatively high, as the savings from RAC activity in the first year are overshadowed by the compliance costs hospitals incur. It falls over time as more savings accrue and hospital compliance costs decrease. Figure 8b plots the MVPF by 2013 under different assumptions. If RAC audits had no deterrence effect and only reclaimed payments, the MVPF would be much larger at 4.55, as each dollar returned to the government's budget would be extremely costly. If RAC audits did not increase hospital administrative costs, the MVPF would be much lower at 1.07 – indicating that RAC audits would be immediately welfare-improving in 2011. The MVPF is also sensitive to assumptions about the effect on patient health: assuming the marginal admission reduces mortality substantially lowers the MVPF, while assuming that it increases mortality increases it (using the upper and lower bound of mortality estimates reported in Currie and Slusky (2020)).

Comparing the MVPF of a revenue-raising policy like RAC audits to the MVPF of an expenditure policy tells us whether combining the two would be welfare-improving. If the two policies have the same distributional incidence and the former is smaller than the latter, then the combined policy is welfare-improving. In this case, the natural policy to combine RAC audits with would be Medicare expenditure itself. Figure 8a plots the MVPF against two MVPFs: 1.63 (the MVPF of Medicare spending, estimated by Finkelstein and McKnight (2008) and Hendren and Sprung-Keyser (2020)), and 1.3 (a commonly-used benchmark for the MVPF (Finkelstein and Hendren, 2020)).[25] The MVPF of RAC audits crosses these thresholds for the MVPF of Medicare expenditure by 2013, and it crosses this threshold for

[24]Patient health may not be the only component of patient welfare that is affected by audits; for example, patients could suffer psychological harm if they are denied admission when they believe it is necessary, but they could also be harmed by an unnecessary admission in the form of wasted time spent in the hospital. However, these other components of patient welfare are difficult to measure and it is unclear what their net effect on patient welfare would be. So, I primarily focus on the effects of a deterred hospitalization on patients' physical health.

[25]See the Policy Impacts Library (available at www.policyimpacts.org) for an extensive database of MVPF estimates on other forms of government expenditure. The revenue raised from RAC audits could instead be spent, for example, on Medicare Part D (MVPF: 1.98) or to subsidize Medicare Advantage plans (MVPF: 1.0). Note that one can only draw welfare conclusions from comparing the MVPF of two programs under the assumption that they have the same distributional incidence.

an MVPF of 1.3 by 2015. Thus, recovering Medicare revenue via RAC audits is welfare improving in the medium- to long-run.

5 Conclusion

The extent to which public programs should monitor third party spending depends on the balance between the money it saves and the social cost of achieving these savings. I study this question in the context of a policy where some of these costs can be observed: monitoring for unnecessary care in Medicare. Here, monitoring saves Medicare money mostly by deterring spending, particularly for unnecessary care. But while these savings did not lead to worsened patient outcomes, they did impose compliance costs on *providers*. In response to increased monitoring, hospitals increased their administrative costs as they invested in technology to detect unnecessary care. But because the compliance costs were primarily incurred upfront and the savings from deterred care accrued over several years, the societal cost per dollar saved by RACs decreases over time. Taken together, the results suggest that Medicare is leaving too much "money on the table," and could stand to monitor more.

More generally, the findings in this paper highlight the importance of accounting for both the direct and *indirect* effects of regulation when evaluating policy. Ignoring the indirect deterrence effects, as policymakers did when assessing the RAC program, can undersell the cost-effectiveness of a policy, leading to insufficient regulation. But failing to consider the indirect *social costs* of regulation, beyond just the direct government cost, may result in suboptimally high levels of oversight. The RAC program serves as an example of how the policy conclusions are sensitive to the inclusion of exclusion of these indirect effects, pointing to the importance of studying them in other policy contexts as well.

References

3M, "3M APCfinderTM Software with Medical Necessity Validation," May 2016.

Agency for Health Care Administration, "Ambulatory and Emergency Department Patient Data Layout," Technical Report 2015.

Alexander, Diane and Molly Schnell, "The Impacts of Physician Payments on Patient Access, Use, and Health," Technical Report w26095, National Bureau of Economic Research July 2019.

Allingham, Michael G. and Agnar Sandmo, "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, November 1972, *1* (3), 323–338.

Almond, Douglas and Joseph J. Doyle, "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays," *American Economic Journal: Economic Policy*, August 2011, *3* (3), 1–34.

American Hospital Association, "Exploring the Impact of the RAC Program on Hospitals Nationwide: Results of AHA RACTrac Survey, 1st Quarter 2012," May 2012.

_ , "Exploring the Impact of the RAC Program on Hospitals Nationwide: Results of AHA RACTrac Survey, 1st Quarter 2014," May 2014.

Andreoni, James, Brian Erard, and Jonathan Feinstein, "Tax Compliance," *Journal of Economic Literature*, 1998, *36* (2), 818–860. Publisher: American Economic Association.

Baron, David P. and David Besanko, "Monitoring, Moral Hazard, Asymmetric Information, and Risk Sharing in Procurement Contracting," *The RAND Journal of Economics*, 1987, *18* (4), 509–532. Publisher: [RAND Corporation, Wiley].

Brot-Goldberg, Zarek, Samantha Burn, Timothy Layton, and Boris Vabson, "Rationing Medicine Through Paperwork: Authorization Restrictions in Medicare," 2021.

Carroll, C., "Impeding Access or Promoting Efficiency? Effects of Rural Hospital Closure on the Cost and Quality of Care," *undefined*, 2019.

Centers for Medicare and Medicaid Services, "Implementation of Recovery Auditing at the Centers for Medicare and Medicaid Services FY2010 Report to Congress," Technical Report 2011.

_ , "Medicare Fee-For-Service 2010 Improper Payment Report," Technical Report 2011.

_ , "Recovery Auditing in the Medicare and Medicaid Programs for Fiscal Year 2011 FY 2011 Report to Congress," Technical Report 2012.

_ , "Hospital Appeals Settlement Process," 2014.

_ , "Clarifying Medical Review of Hospital Claims for Part A Payment," MLN Matters MM10080 May 2017.

—, “2017 Medicare Fee-for-Service Supplemental Improper Payment Data,” Technical Report 2018.

—, “Improper Payment Rates and Additional Data,” Technical Report 2022.

Chan, David C., “Teamwork and Moral Hazard: Evidence from the Emergency Department,” *Journal of Political Economy*, June 2016, *124* (3), 734–770. Publisher: The University of Chicago Press.

Clemens, Jeffrey and Joshua D. Gottlieb, “Do Physicians’ Financial Incentives Affect Medical Treatment and Patient Health?,” *American Economic Review*, April 2014, *104* (4), 1320–1349.

Cooper, Zack, Stuart V Craig, Martin Gaynor, and John Van Reenen, “The Price Ain’t Right? Hospital Prices and Health Spending on the Privately Insured*,” *The Quarterly Journal of Economics*, February 2019, *134* (1), 51–107.

Cubanski, Juliette, Tricia Neuman, and Meredith Freed, “The Facts on Medicare Spending and Financing,” August 2019.

Cunningham, Susan, Mark McMillan, Sara O’Rourke, and Eric Schweikert, “Cracking down on government fraud with data analytics,” October 2018.

Currie, Janet and David Slusky, “Does the Marginal Hospitalization Save Lives? The Case of Respiratory Admissions for the Elderly,” Technical Report w26618, National Bureau of Economic Research January 2020.

Cutler, David, “The Incidence of Adverse Medical Outcomes under Prospective Payment,” *Econometrica*, 1995, *63* (1), 29–50. Publisher: Econometric Society.

Cutler, David M. and Dan P. Ly, “The (Paper)Work of Medicine: Understanding International Medical Costs,” *Journal of Economic Perspectives*, June 2011, *25* (2), 3–25.

Dafny, Leemore S., “How Do Hospitals Respond to Price Changes?,” *American Economic Review*, December 2005, *95* (5), 1525–1547.

de Chaisemartin, C and X D’HaultfŒuille, “Fuzzy Differences-in-Differences,” *The Review of Economic Studies*, April 2018, *85* (2), 999–1028.

Deshpande, Manasi and Yue Li, “Who Is Screened Out? Application Costs and the Targeting of Disability Programs,” *American Economic Journal: Economic Policy*, November 2019, *11* (4), 213–248.

Doyle, Joseph J., John A. Graves, Jonathan Gruber, and Samuel A. Kleiner, “Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns,” *Journal of Political Economy*, February 2015, *123* (1), 170–214. Publisher: The University of Chicago Press.

Dranove, David, Chris Forman, Avi Goldfarb, and Shane Greenstein, "The Trillion Dollar Conundrum: Complementarities and Health Information Technology," *American Economic Journal: Economic Policy*, November 2014, *6* (4), 239–270.

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

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India," *Econometrica*, 2018, *86* (6), 2123–2160. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA12876.

Dunn, Abe, Joshua D. Gottlieb, and Adam Hale Shapiro, "Administration above Administrators: The Changing Technology of Health Care Management," *AEA Papers and Proceedings*, May 2020, *110*, 274–278.

_ , Joshua Gottlieb, Adam Shapiro, and Daniel Sonnenstuhl, "A Denial A Day Keeps the Doctor Away," in "in" April 2021.

Einav, Liran, Amy Finkelstein, and Neale Mahoney, "Provider Incentives and Healthcare Costs: Evidence From Long-Term Care Hospitals," *Econometrica*, 2018, *86* (6), 2161–2219. Publisher: Econometric Society.

Eliason, Paul J., Paul L. E. Grieco, Ryan C. McDevitt, and James W. Roberts, "Strategic Patient Discharge: The Case of Long-Term Care Hospitals," *American Economic Review*, November 2018, *108* (11), 3232–3265.

Ellis, R. P. and T. G. McGuire, "Provider behavior under prospective reimbursement. Cost sharing and supply," *Journal of Health Economics*, June 1986, *5* (2), 129–151.

_ **and** _ , "Hospital response to prospective payment: moral hazard, selection, and practice-style effects," *Journal of Health Economics*, June 1996, *15* (3), 257–277.

Experian Health, "Product Sheet: Medical Necessity," 2022.

Finkelstein, Amy and Matthew J Notowidigdo, "Take-Up and Targeting: Experimental Evidence from SNAP*," *The Quarterly Journal of Economics*, August 2019, *134* (3), 1505–1556.

_ **and Nathaniel Hendren**, "Welfare Analysis Meets Causal Inference," *Journal of Economic Perspectives*, November 2020, *34* (4), 146–167.

_ **and Robin McKnight**, "What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending," *Journal of Public Economics*, 2008, *92* (7), 1644–1668. Publisher: Elsevier.

_ , Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence From Patient Migration*," *The Quarterly Journal of Economics*, November 2016, *131* (4), 1681–1726.

Foster, Anne and B. Scott McBride, "CMS Announces Changes to RAC Program; Temporary Pause in Document Requests," *JD Supra*, March 2014.

Gottlieb, Joshua D., Adam Hale Shapiro, and Abe Dunn, "The Complexity Of Billing And Paying For Physician Care," *Health Affairs*, April 2018, *37* (4), 619–626. Publisher: Health Affairs.

Gowrisankaran, Gautam, Keith A. Joiner, and Jianjing Lin, "How do Hospitals Respond to Payment Incentives?," Working Paper 26455, National Bureau of Economic Research November 2019. Series: Working Paper Series.

Gross, Tal, Adam Sacarny, Maggie Shi, and David Silver, "Regulated Revenues and Hospital Behavior: Evidence from a Medicare Overhaul," 2022.

Gujral, Kritee and Anirban Basu, "Impact of Rural and Urban Hospital Closures on Inpatient Mortality," Working Paper 26182, National Bureau of Economic Research August 2019. Series: Working Paper Series.

Gupta, Atul, "Impacts of Performance Pay for Hospitals: The Readmissions Reduction Program," *American Economic Review*, April 2021, *111* (4), 1241–1283.

Hanna, Rema and Paulina Oliva, "The impact of inspections on plant-level air emissions," *The B.E. journal of economic analysis & policy*, 2010, *10* (1).

Hendren, Nathaniel and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies*," *The Quarterly Journal of Economics*, August 2020, *135* (3), 1209–1318.

Howard, David H. and Ian McCarthy, "Deterrence effects of antifraud and abuse enforcement in health care," *Journal of Health Economics*, January 2021, *75*, 102405.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475. Publisher: [Wiley, Econometric Society].

Jin, Ginger Zhe, Ajin Lee, and Susan Feng Lu, "Medicare Payment to Skilled Nursing Facilities: The Consequences of the Three-Day Rule," Technical Report w25017, National Bureau of Economic Research September 2018.

Kleven, Henrik Jacobsen and Claus Kreiner, "The marginal cost of public funds: Hours of work versus labor force participation," *Journal of Public Economics*, 2006, *90* (10-11), 1955–1973. Publisher: Elsevier.

Kopczuk, Wojciech and Cristian Pop-Eleches, "Electronic filing, tax preparers and participation in the Earned Income Tax Credit," *Journal of Public Economics*, 2007, *91* (7-8), 1351–1367. Publisher: Elsevier.

Krumholz, Harlan M., Kun Wang, Zhenqiu Lin, Kumar Dharmarajan, Leora I. Horwitz, Joseph S. Ross, Elizabeth E. Drye, Susannah M. Bernheim, and Sharon-Lise T. Normand, "Hospital-Readmission Risk — Isolating

Hospital Effects from Patient Effects," *New England Journal of Medicine*, September 2017, *377* (11), 1055–1064. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMsa1702321.

Laffont, Jean-Jacques and Jean Tirole, "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, 1986, *94* (3), 614–641. Publisher: University of Chicago Press.

Leder-Luis, Jetson, "Can Whistleblowers Root Out Public Expenditure Fraud? Evidence from Medicare," July 2020, p. 60.

Lopez, Eric, Gretchen Jacobson, Tricia Neuman, and Larry Levitt, "How Much More Than Medicare Do Private Insurers Pay? A Review of the Literature," April 2020.

Magat, Wesley A. and W. Kip Viscusi, "Effectiveness of the EPA's Regulatory Enforcement: The Case of Industrial Effluent Standards," *The Journal of Law and Economics*, October 1990, *33* (2), 331–360. Publisher: The University of Chicago Press.

Meckel, Katherine, "Is the Cure Worse Than the Disease? Unintended Effects of Payment Reform in a Quantity-Based Transfer Program," *American Economic Review*, June 2020, *110* (6), 1821–1865.

Medicare Contractor Management Group, "Medicare Administrative Contractor Workload Transition Handbook," October 2017.

Medicare Payment Advisory Commission, "Hospital Short-Stay Policy Issues," Technical Report June 2015.

_ , "July 2020 Data Book: Health Care Spending and the Medicare Program," Technical Report July 2020.

Nicholas, Lauren Hersch, Caroline Hanson, Jodi B. Segal, and Matthew D. Eisenberg, "Association Between Treatment by Fraud and Abuse Perpetrators and Health Outcomes Among Medicare Beneficiaries," *JAMA internal medicine*, January 2020, *180* (1), 62–69.

Office of the Inspector General, "Medicare Improperly Paid Acute-Care Hospitals $54.4 Million for Inpatient Claims Subject to the Post-Acute-Care Transfer Policy," Technical Report A-09-19-03007 November 2019.

Olken, Benjamin A., "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, April 2007, *115* (2), 200–249. Publisher: The University of Chicago Press.

Powell, W. Ryan, Farah A. Kaiksow, Amy J. H. Kind, and Ann M. Sheehy, "What Is an Observation Stay? Evaluating the Use of Hospital Observation Stays in Medicare," *Journal of the American Geriatrics Society*, 2020, *68* (7), 1568–1572. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jgs.16441.

ResDAC, "Anomaly in MedPAR/inpatient claims when counting emergency department visits," May 2022.

Roberts, James, Ryan McDevitt, Paul Eliason, Jetson Leder-Luis, and Riley League, "Enforcement and Deterrence of Medicare Fraud: The Case of Non-emergent Ambulance Rides," in "in" 2021.

Rose, Liam, "The Effects of Skilled Nursing Facility Care: Regression Discontinuity Evidence from Medicare," *American Journal of Health Economics*, January 2020, *6* (1), 39–71. Publisher: The University of Chicago Press.

Sacarny, Adam, "Adoption and learning across hospitals: The case of a revenue-generating practice," *Journal of Health Economics*, July 2018, *60*, 142–164.

Sheehy, Ann M., Fangfang Shi, and Amy J.H. Kind, "Identifying Observation Stays In Medicare Data: Policy Implications of a Definition," *Journal of hospital medicine*, February 2019, *14* (2), 96–100.

Silverman, Elaine and Jonathan Skinner, "Medicare upcoding and hospital ownership," *Journal of Health Economics*, March 2004, *23* (2), 369–389.

Skinner, Jonathan, "Causes and Consequences of Regional Variations in Health Care," Handbook of Health Economics, Elsevier 2011.

Slemrod, Joel and Shlomo Yitzhaki, "Integrating Expenditure and Tax Decisions: The Marginal Cost of Funds and the Marginal Benefit of Projects," *National Tax Journal*, 2001, *54* (2), 189–202. Publisher: National Tax Association.

_ and _ , "Tax avoidance, evasion, and administration," Handbook of Public Economics, Elsevier 2002.

Span, Paula, "In the Hospital, but Not Really a Patient," June 2012. Cad: 0.

US Department of Health and Human Services Office of Inspector General, "Hospitals' Use of Observation Stays and Short Inpatient Stays for Medicare Beneficiaries," Technical Report OEI-02-12-00040 July 2013.

U.S. Government Accountability Office, "Federal Government Contracting for Fiscal Year 2018," May 2019.

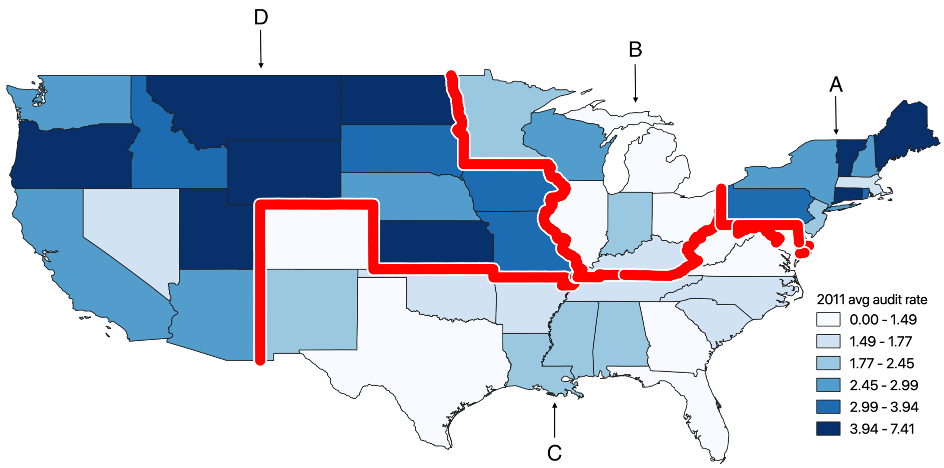
Venkatesh, Arjun K., Hao Mei, Keith E. Kocher, Michael Granovsky, Ziad Obermeyer, Erica S. Spatz, Craig Rothenberg, Harlan M. Krumholz, and Zhenqui Lin, "Identification of Emergency Department Visits in Medicare Administrative Claims: Approaches and Implications," *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, April 2017, *24* (4), 422–431.

Zwick, Eric, "The Costs of Corporate Tax Complexity," *American Economic Journal: Economic Policy*, May 2021, *13* (2), 467–500.

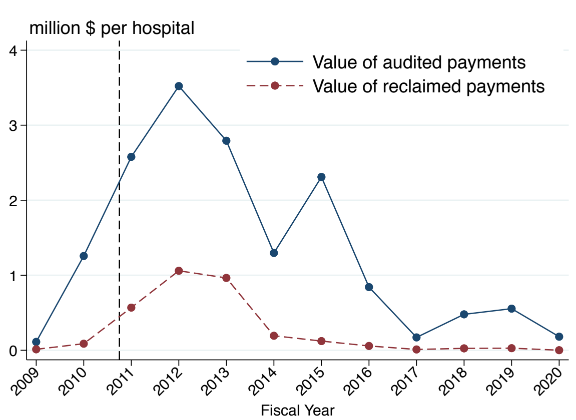
6 Figures

Figure 1. RAC Audit Activity

(a) Average 2011 Hospital Audit Rates by State and RAC Regions

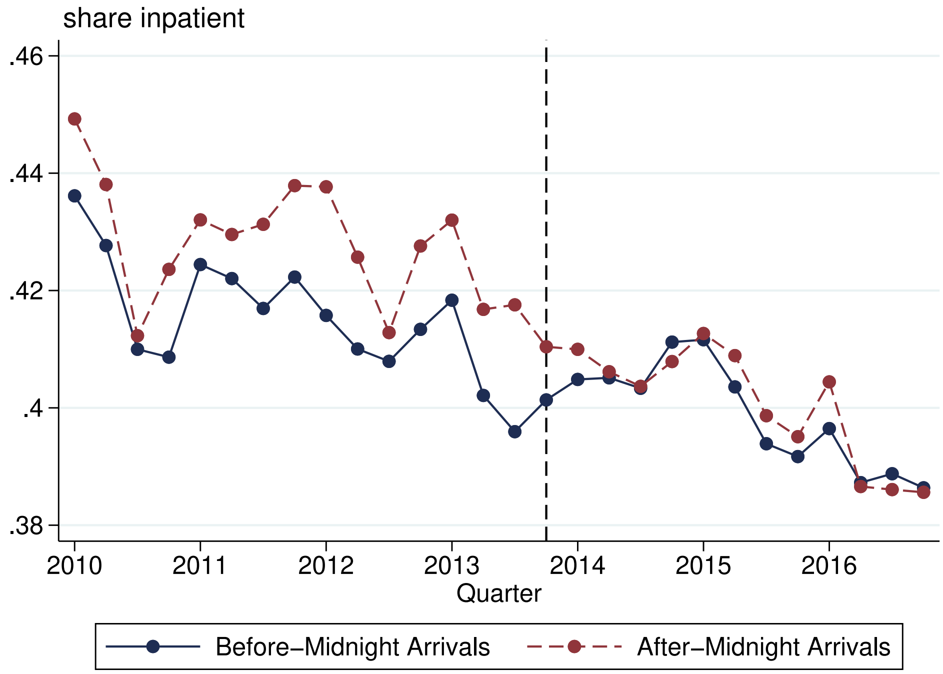


(b) Value of Audited Inpatient Payments and Net Reclaimed Payments per Hospital, by Year of Audit



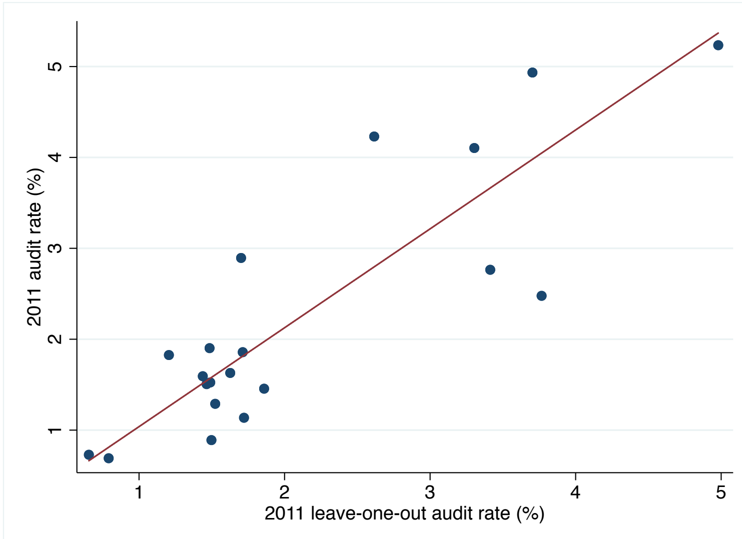
Panel (a) plots the 2011 average state audit rates, where audit rate is defined as the share of a hospital's 2008-2011 claims that were audited by RACs. The RAC regions are Region A (Northeast), Region B (Midwest), Region C (South), and Region D (West). Darker shades denote a higher audit rate. The red line demarcates RAC regions. Panel (b) plots the average per-hospital value of inpatient payments audited by RACs and the net reclaimed payments, by year of audit. Net reclaimed payments are defined as the sum of reclaimed payments from overpayments minus refunded payments from underpayments. These values are based on RACs' original reclaimed or refunded payments at the time of audit. Data: MEDPAR claims and CMS audit data.

Figure 2. Inpatient Admission Rates from ED, Before vs. After-Midnight ED Arrivals in Florida



This figure plots the share of traditional Medicare patients admitted as inpatient from the emergency department, among Florida patients who arrived within three hours before midnight (9:00-11:59PM), in the blue solid line, and three hours after midnight (12:00-2:59AM), in the red dashed line. The dashed vertical line denotes 2013Q3, which is when the Two Midnights rule is implemented. Data: HCUP SID/SEDD.

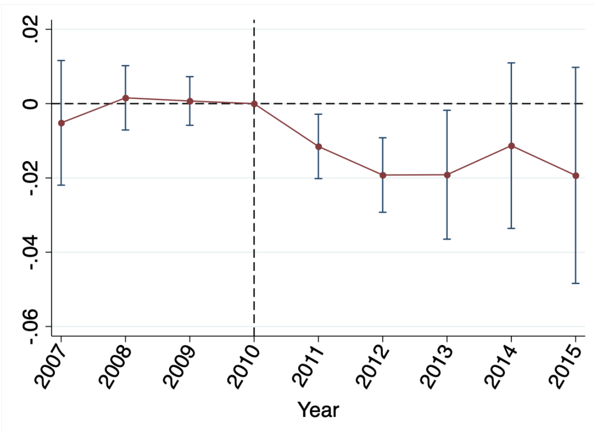
Figure 3. Binscatter of 2011 Leave-One-Out State Audit Rate and 2011 Hospital Audit Rate, Border Hospital Sample



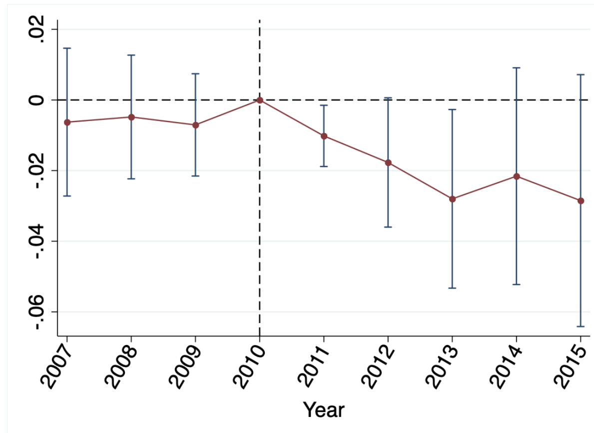
This figure plots a binscatter of the 2011 hospital audit rate compared to the 2011 leave-one-out state audit rate. The 2011 audit rate is defined as the share of 2008-2011 inpatient claims that were audited by RACs in 2011. The leave-one-out state audit rate is defined as the average audit rate of all other hospitals in the same state as a given hospital. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group. Data: MEDPAR claims and CMS audit data.

Figure 4. Event Studies on Effect of 2011 Audit Rate on Hospital Outcomes

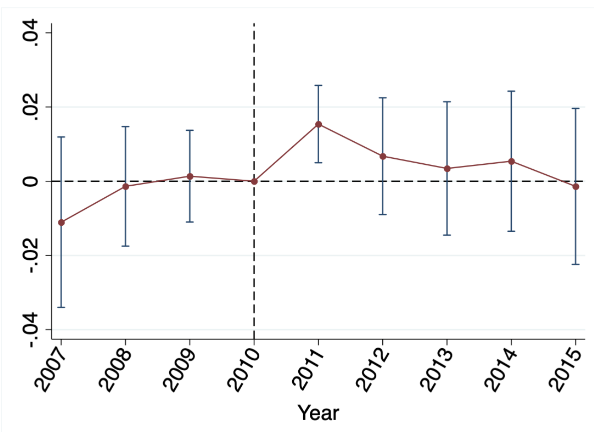
(a) Log Medicare admissions



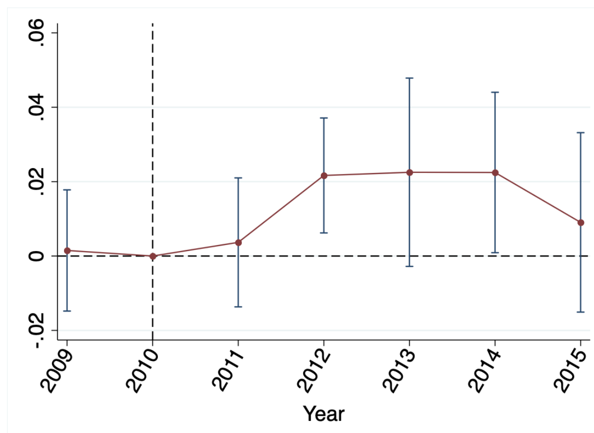
(b) Log Medicare inpatient revenue



(c) Log hospital administrative costs



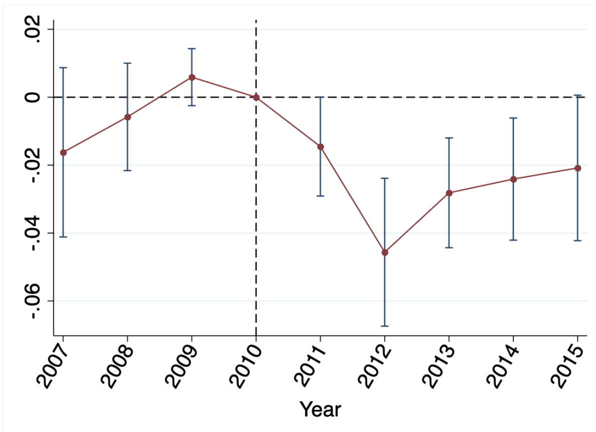
(d) Indicator for installing medical necessity software



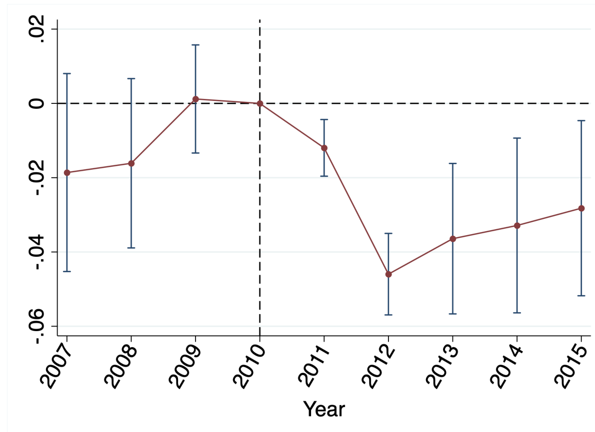
This figure plots event studies of the reduced form coefficients and 95% confidence interval in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome. Medicare admissions and revenue are from MEDPAR. Inpatient revenue is the sum of all Medicare inpatient payments. Net administrative costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Indicator for installing software is equal to 1 if a hospital reports the status of a medical necessity software as "contracted/not yet installed," "installation in process," and "to be replaced" in HIMSS. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

Figure 5. Event Studies on Effect of 2011 Audit Rate on Medicare Admissions and Revenue, by Length of Stay

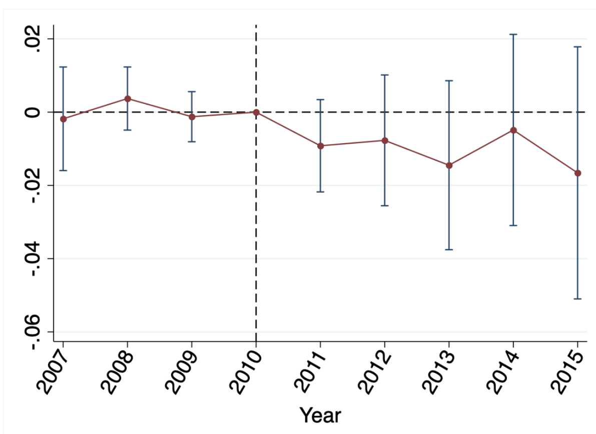
(a) Log Medicare admissions, LOS ≤ 2



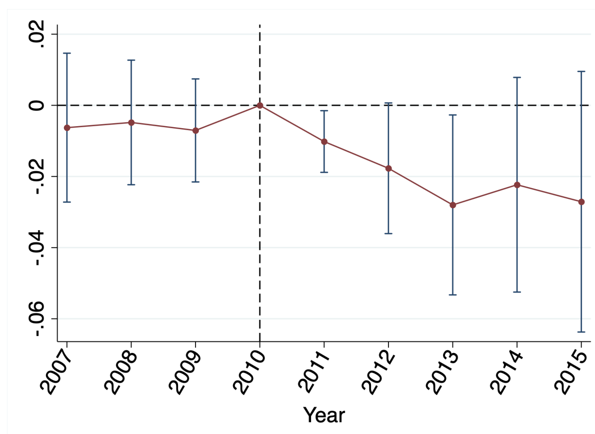
(b) Log Medicare inpatient revenue, LOS ≤ 2



(c) Log Medicare admissions, LOS > 2

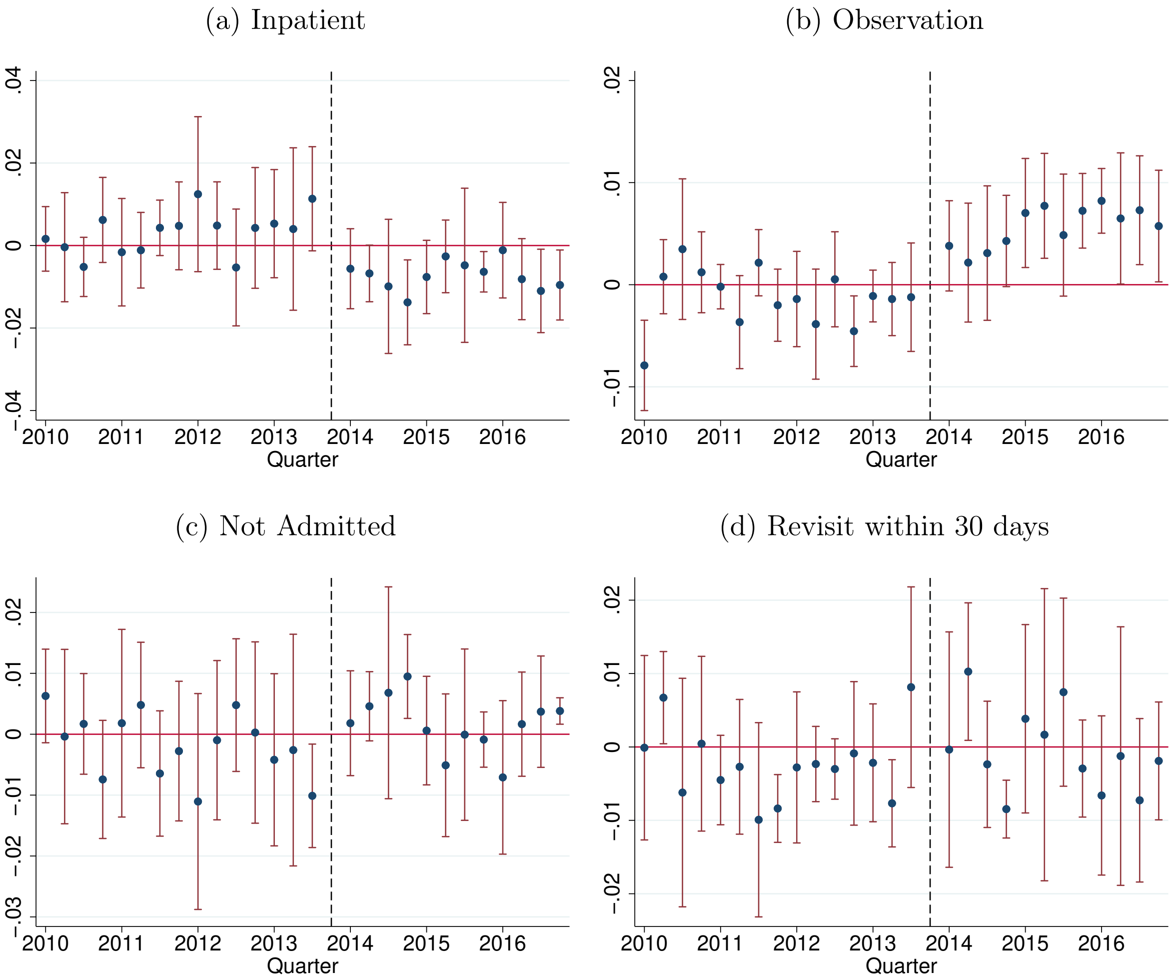


(d) Log Medicare inpatient revenue, LOS > 2



This figure plots event studies of the reduced form coefficients and 95% confidence interval in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome. Medicare volume and revenue of short stay admissions and longer admissions are from MEDPAR. Length of stay is counted as the difference in days between the admission and discharge date. Inpatient revenue is the sum of all Medicare inpatient payments. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

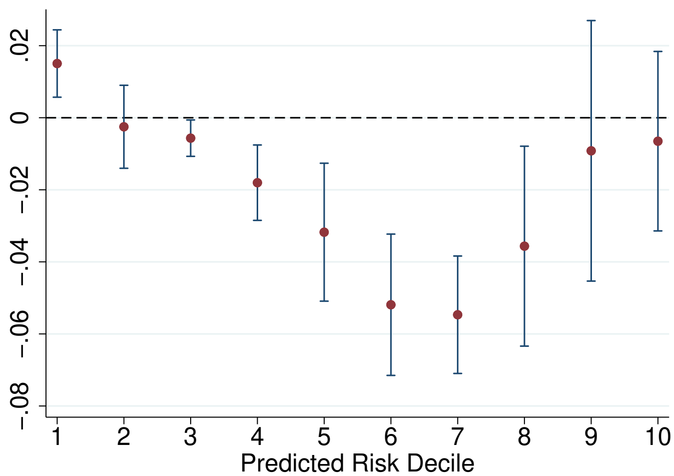
Figure 6. Event Studies on Effect of After-Midnight ED Arrival on Patient Status and Outcomes



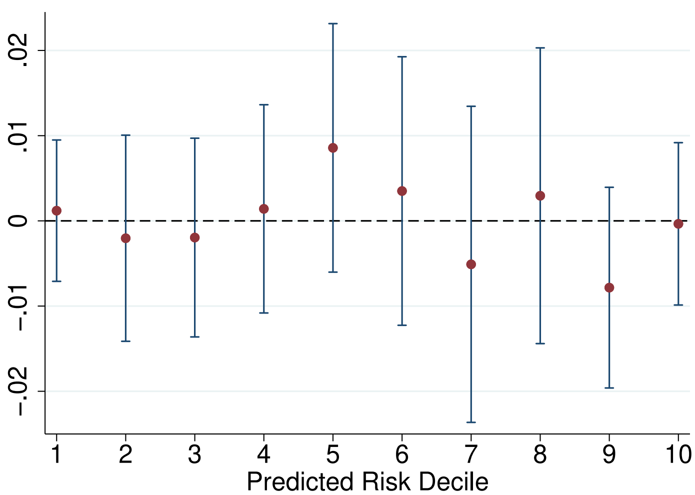
This figure plots the coefficients and 95% confidence intervals for β^τ on $\mathbb{1}[q=\tau]\times\mathbb{1}[T_v \geq 00{:}00]$ of the specification in Equation 7, where $\mathbb{1}[q=\tau]$ is an indicator for whether the visit occurred in quarter τ, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. The results are clustered at the ED arrival hour and quarter level. The omitted quarter is 2013Q3. "Inpatient" is an indicator for whether the patient was eventually admitted as inpatient from the ED. "Observation" is an indicator for whether the patient was placed in observation status and was never admitted. "Not Admitted" is an indicator equal to one when a patient is neither admitted nor placed in observation status. "Revisit within 30 days" is an indicator for whether the patient had another ED visit or inpatient stay within 30 days of the ED visit. Sample consists of traditional Medicare patients who arrived in the ED within 3 hours of midnight in a Florida hospital. Regression includes hospital, hospital-quarter, hospital-hour fixed effects, and controls for age-sex bin, race, Hispanic indicator, point of origin indicator, last ED visit within 30 days indicator, number of chronic conditions, and zip code income. Data: HCUP SID/SEDD.

Figure 7. Heterogeneity of After-Midnight ED Arrival Coefficient by Patient Severity

(a) Inpatient



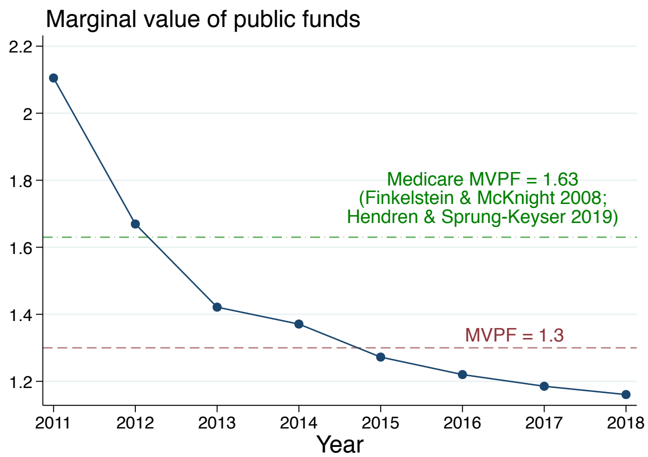
(b) Revisit in 30 days



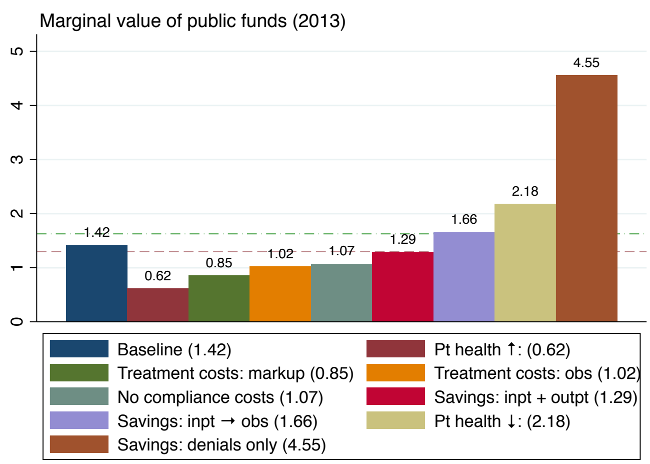
This figure plots estimates and 95% confidence intervals of the β coefficient in Equation 7, interacted with an indicator for predicted severity decile. β is the coefficient on $\mathbb{1}[q \geq 2013Q3] \times \mathbb{1}[T_v \geq 00{:}00]$, where $\mathbb{1}[q \geq 2013Q3]$ is an indicator for whether the visit occurred after 2013Q3, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. The top panel plots results for an indicator for whether the patient was admitted as inpatient from the ED, and the bottom panel plots results for an indicator for whether the patient revisited any hospital in Florida within 30 days of the ED visit. The results are clustered at the ED arrival hour and quarter level. Patient risk is predicted by estimating a logit using ED visits between 9:00AM and 3:00PM of an indicator for being admitted within 30 days of an ED visit on patient demographics, current ED visit information, and information on any prior visits in the last 365 days. Demographics include age-bin, sex, race, Hispanic indicator, point of origin indicator, and mean zip code income. Information on current visit includes hospital and quarter. Information on previous visits includes the number of visits/inpatient stays/length of stay in the last month or last year, as well as any diagnoses and procedures recorded in stays within the last month or last year. Figure E19 plots the mean outcomes for each decile. Data: HCUP SID/SEDD.

Figure 8. Marginal Value of Public Funds Calculation

(a) MVPF By Year, Baseline Calculation



(b) MVPF By 2013 Under Different Assumptions



This figure plots the marginal value of public funds (MVPF) of an increase in the 2011 RAC audit rate. Panel (a) plots the MVPF of savings between 2011 and a given year under the baseline assumptions, compared to MVPF values of 1.3 and 1.63 (Finkelstein and McKnight, 2008; Hendren and Sprung-Keyser, 2020). Panel (b) plots the MVPF by 2013 under different assumptions detailed in Appendix Section D.2. "Patient health ↑:" and "Patient health ↓:" assume that patient mortality decreases or increases from the marginal admission, respectively. "Treatment costs: markup" assumes that treatment costs are a constant markup of inpatient payments, and "Treatment costs: obs" assumes that treatment costs are a constant markup of observation stay payments. "No compliance costs" assumes there is no increase in hospital admin costs. "Savings: inpt + outpt" uses estimates on the effects of inpatient and outpatient revenue, "Savings: inpt → obs" assumes all deterred inpatient stays become observation stays, and "Savings: denials only" assumes there is no deterrence effect.

7 Tables

Table 1. Hospital Summary Statistics by RAC Region

	(1)	(2)	(3)	(4)
	RAC Region			
	A	*B*	*C*	*D*
	A. Hospital Characteristics			
2011 audit rate	3.01	1.79	1.36	3.33
	(2.29)	(1.21)	(1.18)	(2.73)
Share urban	0.83	0.70	0.64	0.82
Share non-profit	0.88	0.79	0.46	0.63
	(0.32)	(0.41)	(0.50)	(0.48)
Share for-profit	0.05	0.09	0.29	0.19
Share government	0.07	0.12	0.24	0.18
Beds	238.22	198.04	194.41	193.59
	(194.54)	(170.28)	(186.64)	(146.62)
Total cost (million \$)	271.89	211.01	154.97	218.05
Net admin costs (million \$)	36.00	33.38	22.24	33.47
	B. Medicare Inpatient Admission Characteristics			
Admissions	4264.70	3845.22	3262.61	2928.68
	(3591.67)	(3383.92)	(3260.47)	(2399.90)
Mean payment (\$)	9349.37	8177.97	7578.76	10393.64
	(3461.79)	(2433.87)	(2663.76)	(3501.44)
Total payments (million \$)	45.75	36.03	29.15	32.65
	(53.88)	(40.65)	(35.72)	(32.25)
Average short stay share	0.28	0.32	0.31	0.33
	(0.07)	(0.07)	(0.08)	(0.07)
Average circulatory diagnosis share	0.23	0.22	0.21	0.20
	(0.07)	(0.07)	(0.08)	(0.08)
Observations	489	571	1237	663
N Border Hospitals	41	184	191	94

This table presents 2010 summary statistics of hospital characteristics and Medicare inpatient admissions by RAC region. Standard deviation is in parentheses. Bed size, urban status, and profit type status come from the Medicare Provider of Services file. Total and administrative costs come from HCRIS. Medicare admissions and inpatient stay characteristics are from MEDPAR. Mean inpatient characteristics are defined as the average of each hospital's average (i.e., weighted by hospitals rather than claims). Short stay share is the share of Medicare admissions with length of stay ≤ 2. Circulatory diagnosis share is the share of Medicare admissions with a circulatory Major Diagnostic Category diagnosis. The border sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

Table 2. Patient Summary Statistics by ED Arrival Hour

	(1)	(2)
	ED Arrival Hour	
	Before MN	*After MN*
Share inpatient	0.40 (0.49)	0.42 (0.49)
Share observation	0.05 (0.21)	0.05 (0.22)
Average charges ($)	23966.55 (43649.05)	25881.27 (50655.54)
Average age	68.04 (17.33)	68.22 (17.28)
Share white	0.78 (0.41)	0.77 (0.42)
Share hispanic	0.12 (0.32)	0.11 (0.31)
Share female	0.57 (0.50)	0.54 (0.50)
Average n of chronic conditions	3.95 (3.57)	4.17 (3.64)
Share inpatient in last 30 days	0.13 (0.33)	0.14 (0.34)
Share hospital visit in last 30 days	0.28 (0.45)	0.30 (0.46)
Share hospital visit in next 30 days	0.27 (0.45)	0.29 (0.45)
Share hospital visit in next 60 days	0.38 (0.48)	0.39 (0.49)
Share hospital visit in next 90 days	0.44 (0.50)	0.45 (0.50)
Observations	32793	18467

This table presents summary statistics of characteristics of traditional Medicare patients in Florida who arrived in the ED within 3 hours of midnight in 2013Q2. Standard deviation is in parentheses. "Share inpatient" is the share of ED patients admitted to inpatient (this includes patients who could have initially been placed in observation and eventually admitted). "Share observation" is the share of patients who are placed in outpatient observation only. Data: HCUP SID/SEDD.

Table 3. Event Studies of Effect of 2011 Audit Rate on Hospital Outcomes, 2011-2015 Coefficients

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall		LOS $\leq$ 2		Admin Costs	Software Installation
	Log Adm.	*Log Rev.*	*Log Adm.*	*Log Rev.*	*Log Costs*	*Medical Necc.*
2011 audit rate × 2011	-0.0115** (0.0044)	-0.0102** (0.0044)	-0.0145* (0.0074)	-0.0120*** (0.0039)	0.0154*** (0.0053)	0.0037 (0.0088)
2011 audit rate × 2012	-0.0192*** (0.0051)	-0.0177* (0.0093)	-0.0457*** (0.0111)	-0.0460*** (0.0056)	0.0068 (0.0080)	0.0217** (0.0079)
2011 audit rate × 2013	-0.0191** (0.0089)	-0.0280** (0.0129)	-0.0282*** (0.0082)	-0.0364*** (0.0103)	0.0034 (0.0092)	0.0225* (0.0129)
2011 audit rate × 2014	-0.0113 (0.0114)	-0.0216 (0.0157)	-0.0241** (0.0092)	-0.0329** (0.0120)	0.0054 (0.0096)	0.0225* (0.0110)
2011 audit rate × 2015	-0.0193 (0.0148)	-0.0285 (0.0182)	-0.0208* (0.0109)	-0.0282** (0.0107)	-0.0014 (0.0107)	0.0090 (0.0123)
Hosp FE	X	X	X	X	X	X
Nbr group FE	X	X	X	X	X	X
N Hosp	510	510	510	510	510	506
Obs	52139	52139	52139	52118	52107	36906
F	12.5	12.5	12.5	13.36	12.45	13.87

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors are in parentheses and are clustered at the state and border segment level. This table reports the coefficients of the reduced form event study in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome. For brevity, the pre-2011 coefficients are estimated but not reported in the table. Omitted year is 2010. Columns 1 and 2 report the effect on the log number of Medicare inpatient admissions and log Medicare inpatient revenue from the MEDPAR data, and columns 3 and 4 report the effect on short stay admissions and revenue. Column 5 reports the effect on log net administrative costs from HCRIS data. Net administrative costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Column 6 reports the effect on an indicator for installing medical necessity software application, which is equal to 1 if a hospital reports the status of a medical necessity software as "contracted/not yet installed," "installation in process," and "to be replaced" in the HIMSS data. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

Table 4. After-Midnight ED Arrival Hour Difference-in-Difference Coefficients on Patient Status and Revisits

	(1)	(2)	(3)	(4)	(5)
	Medicare				Non-Medicare
	Inpatient	*Observation*	*Not Admitted*	*Revisit 30d*	*Inpatient*
β	-0.007***	0.007***	0.000	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Pre-reform mean	0.420	0.042	0.538	0.259	0.126
Estimate as % of mean	1.67	16.67	0.00	0.39	0.79
Observations	1254857	1254857	1254857	1254857	7428583

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors are in parentheses and are clustered at the ED arrival hour and quarter level. This table reports the β coefficient on $\mathbb{1}[q \geq 2013Q3] \times \mathbb{1}[T_v \geq 00{:}00]$ of the specification in Equation 7, where $\mathbb{1}[q \geq 2013Q3]$ is an indicator for whether the visit occurred after the Two Midnights rule was implemented in 2013Q3, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. "Inpatient" is an indicator for whether the patient was eventually admitted as inpatient from the ED. "Observation" is an indicator for whether the patient was placed in observation status and was never admitted. "Not Admitted" is an indicator equal to one when a patient is neither admitted nor placed in observation status. "Revisit within 30 days" is an indicator for whether the patient had another ED visit or inpatient stay within 30 days of the ED visit. Sample for columns 1-4 consists of traditional Medicare patients who arrived in the ED within 3 hours of midnight in a Florida hospital. The sample for column 5 consists of all non-Medicare patients who arrived in the ED within 3 hours of midnight in a Florida hospital. Regression includes hospital, hospital-quarter, hospital-hour fixed effects, and controls for age-sex bin, race, Hispanic indicator, point of origin indicator, last ED visit within 30 days indicator, number of chronic conditions, and zip code income. Data: HCUP SID/SIDD.

Table 5. Marginal Value of Public Funds Baseline Parameters

	A. Estimates
Effect on admissions	2011-2015: estimates after 2015: 2015 estimate
Effect on compliance costs	2011-2015: estimates after 2015: 0
Payments demanded	2011-2015: estimates after 2015: 0
2011-2015: estimates after 2015: 0	
Avg 2010 inpatient revenue	$15,029,306
Avg 2010 compliance cost	$12,822,887
	B. Parameters
RAC contingency fee	10.75%
Marginal value of public funds	1.3
Discount rate	2%
Share of demanded pmts refunded	68%

This table lists the parameters and assumptions for the MVPF calculation depicted in Figure 8a. Effects on admissions and compliance costs are from Table 3. Payments demanded are from Figure E11. The 2010 hospital revenue and hospital compliance costs are the median values for hospitals in the border hospital sample.

A Additional Policy Context

A.1 Medicare Inpatient Prospective Payment System and Short Stays

Medicare pays for inpatient hospital admissions through the inpatient prospective payment system (IPPS), in which Medicare pays a fixed amount per inpatient stay within broad categories of diagnoses called Medicare Severity Diagnosis Related Groups (MS-DRGs, also referred to as DRGs). The prospective payment system was introduced in 1983 with the intent of incentivizing providers to reduce healthcare costs (Ellis and McGuire, 1986). Hospitals keep the difference between the DRG payment and the costs to treat the patient, so they have an incentive to keep costs low. The payment rate for each DRG reflects the national average cost of treating a patient across all cases, and it is revised each year based on claims data in the last two years. The per-stay payment is adjusted based on a patient's pre-existing chronic conditions in order to account for the patient's diagnosis severity. It is also adjusted by hospital-specific factors such as a hospital's wage index, teaching status, share of low-income patients, and number of unusually costly outlier cases. The prospective payment system generally works well to keep inpatient hospital spending relatively low for the Medicare program (Lopez et al., 2020).

However, one persistent issue with IPPS that has been noted by policymakers is the high number of short stays. A CMS report found that "a large percentage of medically unnecessary [payment] errors are related to hospital stays of short duration... these services should have been rendered at a lower level of care" (Centers for Medicare and Medicaid Services, 2011b). One less intensive alternative to an inpatient stay is an outpatient observation stay, which consists of short-term (often diagnostic) services provided at the hospital while a physician decides whether to formally admit a patient as inpatient or send them home. Observation stays typically last less than forty-eight hours and are billed as an outpatient service (Medicare Payment Advisory Commission, 2015).

From the patient's point of view, it is often difficult to differentiate between an observation stay and a short inpatient stay (Span, 2012). Thus, a hospital's costs for an observation stay are likely similar to the costs for a short inpatient stay. However, hospitals earn much more from Medicare for admitting a patient for a short inpatient stay rather than for an outpatient observation stay: among DRGs common to both inpatient and observation stays, Medicare payments for inpatient stays were two to three times higher than payments for observation stays (Medicare Payment Advisory Commission, 2015).

Policymakers considered various alternative solutions to address unnecessary short stays before settling on RAC audits. They were wary of reducing the payment rate for short stays or penalizing high rates of short stays, due to concerns that hospitals would simply keep patients for longer to evade these policies (Medicare Payment Advisory Commission, 2015). There is evidence that hospitals delay discharging patients if they have an incentive to do so (Jin et al., 2018). Additionally, short stays constitute almost a third of inpatient stays; their prevalence suggests that not all short stays are unnecessary, and cutting payments for short stays across the board would reduce payments for some necessary stays.

A.2 RAC Program Details

RAC Regions In the context of medical claims processing and reviews, the jurisdictions used for RAC regions are unique. Medicare Administrative Contractors (MACs) are contractors who process Medicare claims *before* payment; they operate in different, smaller regions than RAC regions. The RAC regions do align with the regions of Durable Medical Equipment MACs. However, they only process payments for durable medical equipment like prosthetics, orthotics, and other devices, and they do not process claims for medical services (Medicare Contractor Management Group, 2017). To hire RAC firms for each region, Medicare posts a separate contract solicitation for each region, and firms submit separate bids.

RAC Firms The four firms originally contracted to conduct RAC audits in 2010 were Health Data Insight, Cotiviti, CGI, and Performant Recovery (Centers for Medicare and Medicaid Services, 2011a). Some firms focus on healthcare (for example, Health Data Insight, Cotiviti), while others serve other government agencies and corporations as well (for example, CGI, Performant Recovery). Other clients of the RAC firms include state tax authorities, student loan companies, private health insurance companies, the Internal Revenue Service, the National Health Service in the UK, and Public Health England.

RAC Audit Process RACs conduct postpayment reviews to identify and correct overpayments or underpayments for claims for inpatient care, outpatient care, long-term care, and durable medical equipment in the last three years. Figure E1 illustrates the claims auditing and appeal process, using 2011 inpatient audits as an example. Each RAC develops and runs its own proprietary algorithm on claims data to identify claims with potential payment errors. In 2011, RACs' auditing scope for inpatient claims included incorrect or incomplete coding, DRG validation, and medical necessity reviews. Five percent of audits were "automated reviews," which rely solely on claims data to make a determination based on clearly outlined Medicare policies. The rest of the audits were "complex reviews," in which a medical professional (for example, coder, nurse, or therapist) employed by the RAC submits a medical record request and manually reviews all documentation associated with an inpatient stay. It is up to the medical professional to determine whether an overpayment or underpayment was made. Once the complex review is finished, RACs send a demand letter to providers that outlines whether a payment error was identified, the amount of overpayment or underpayment demanded, and references supporting the decision. Fifty-seven percent of complex reviews in 2011 resulted in no finding, 37 percent resulted in an overpayment demand (in which providers must return payment back to Medicare), and 6 percent resulted in an underpayment demand (in which Medicare returns payment to the provider). Providers can appeal demands by first requesting a redetermination by the RAC and then escalating it to higher levels of appeals – for example, by requesting that a separate contractor reconsider the case, requesting a hearing by an administrative law judge, or escalating it to a review by the Medicare Appeals Council.

Timeline of the RAC Program The RAC program was first proposed as part of the Medicare Modernization Act of 2003. After an initial pilot demonstration from 2005 to 2008 in select states, the RAC program was implemented nationally in 2010 (Centers for Medicare and Medicaid Services, 2011a). At first, RACs were authorized only to audit claims with complex coding issues and for DRG validation. Each year, Medicare expanded the scope of RAC audits, and in 2011 it expanded the scope to include medical necessity reviews of inpatient claims (Centers for Medicare and Medicaid Services, 2012). As shown in Figure 1b, RAC audit activity peaked in 2011–13, then dropped precipitously in 2014. The peak corresponds with the period in which RACs were authorized to audit inpatient claims for medical necessity.

In the face of a sudden rise in auditing and overpayment demands, hospitals began mounting a campaign to fight back. Hospitals started appealing high volumes of RAC determinations, and some hospital systems worked with the American Hospital Association (AHA) to file lawsuits and complaints against Medicare over RAC audits.[26] Between 2011 and 2013, the number of appeals that reached the administrative-judge level of the appeals process increased by 500 percent, and by mid-2014 there was a backlog of eight hundred thousand appeals at that level (Medicare Payment Advisory Commission, 2015). The AHA also began tracking the effect of RAC activity on its own through the quarterly RACTrac Survey of hospitals. Many hospitals reported that RAC audits imposed significant administrative burdens on them; for example, 11 percent of hospitals reported costs associated with managing the RAC program of over $100,000 (American Hospital Association, 2014).

Hospitals and industry stakeholders filed several complaints with Medicare stating that RAC audits were overly aggressive. As a result, in 2014 Medicare paused almost all RAC audits by significantly limiting their scope (Foster and McBride, 2014). Other Medicare contractors such as MACs picked up additional review responsibilities after the RAC audits were paused.[27] Medicare maintained that the pause on RAC audits was temporary and

[26] See the AHA website for a list of all past and ongoing litigation: https://www.aha.org/legal/past-litigation (link).

[27] For example, MACs conducted a program called "Teach, Probe, and Educate" in which they targeted

would resume at previous levels, but it is clear from Figure 1b that RAC auditing never returned to its peak level after the pause. The pause began at the end of 2014Q1 and was originally meant to end in 2014Q3. After several quarters of delayed resumption, inpatient RAC audits finally resumed in 2015Q4, although they were subject to limitations to reduce the administrative burden on providers. In August 2014, Medicare announced a one-time option to settle appeals by offering hospitals 68 percent of each appealed denied inpatient claim, in exchange for hospitals dropping all of their appeals rather than settling them one by one. As a result, hospitals dropped almost 350,000 appeals in exchange for $1.5 billion in settled denials (Centers for Medicare and Medicaid Services, 2014).

A.3 Characteristics of Audits and Audited Hospitals

Given Medicare policymakers' focus on short stays as the main source of unnecessary admissions, I examine audit frequency as a function of an admission's length of stay in Figure E2. Admissions with a length of stay of two or fewer days have much higher rates of auditing than longer admissions. Admissions with a length of stay fewer than two days have an average audit rate of 4.2 percent, while admissions with a length of stay more than two days have an average audit rate of 0.7 percent. The majority of audits of short stays result in the full payment being reclaimed (Figure E3). I also consider audit frequencies by diagnosis – circulatory diagnoses are subject to more audits relative to other diagnosis types.

I next consider hospital-level characteristics and their correlation with audit rate in Figure E4. The RAC region a hospital is in is highly correlated with its audit rate. Within each region, rural hospitals, small hospitals, non-profit hospitals, and hospitals with a higher share of short stay Medicare admissions are more likely to be audited.

Although almost every hospital was subject to an audit by 2020, in any given year there is a substantial portion of hospitals that do not face any audits. In 2011, 15 percent of hospitals had an audit rate of 0 percent. The share of hospitals with no audits varies across RAC regions from 2 to 23 percent (Figure E7).

B Robustness and Placebo Tests

Hospital-Level Analysis As a robustness test, in Figure E14 I regress on a hospital's denial rate – the share of claims for which a denial is made after audit – rather than its audit rate. Equation 9 defines the relationship between denial rate and audit rate.

hospitals with high payment errors and conducted education sessions. If hospitals failed to improve their payment accuracy sufficiently after three rounds of education sessions, then they were referred to Medicare for further remediation.

$$\textit{Denial Rate}_{ht} = \underbrace{P(\textit{Audit})_{ht}}_{\textit{Audit Rate}} \times \underbrace{P(\textit{Demand}|\textit{Audit})_{ht}}_{\textit{Demand Rate}} \tag{9}$$

Since 41 percent of audits in 2011 resulted in a demand in the main sample, one would expect that a hospital's response to a one-percentage point increase in the denial rate should be about twice the response to one percentage point increase in the audit rate. Indeed, this is what the results reflect; for example, hospitals reduced admissions by 2.5 percent in 2012 in response to a one-percentage point increase in the 2011 audit rate, and they reduced admissions by 5.7 percent in 2012 in response to a one-percentage point increase in the denial rate.

In Figure E15, I show that the results are robust to alternative sample definitions. Figure E15a reproduces the event study from the main specification for the outcome of log Medicare admissions, in which the sample is defined as all hospitals within 100 miles of the RAC border and the coefficient is scaled by the correlation between a hospital's audit rate and its leave-one-out state audit rate. This is robust to changing the sample to all hospitals within 50 miles (Figure E15b) or 150 miles (Figure E15c) of the border, although the results are noisier with a shorter distance. One concern with boundary discontinuity identification strategies is the potential for spillovers among hospitals very close to the border. For example, if patients were redirected from a hospital near the border in a high-audit rate state to a nearby hospital in a low-audit rate state, then this would bias the coefficients upward. Figure E15d shows similar results when restricting the sample to hospitals that are at least 10 miles away from the border, demonstrating that the result is not driven by such spillovers. Finally, Figure E15e shows that the results are similar when restricting the sample to hospitals with audit rates greater than 0 percent, meaning that the results are driven by variation in auditing across hospitals on the intensive, rather than the extensive, margin.

Figure E16 shows that the results are robust to using alternative instruments to scale the reduced form effect. The main specification instruments for a hospital's audit rate using the leave-one-out state audit rate in order to capture the variation in audit intensity that is unrelated to the hospital's own behavior. Figure E16a plots the results of using the state audit rate (which includes the hospital) as an instrument. Figure E16c shows that the results using the leave-one-out RAC region audit rate, rather than the state audit rate, are similar.

While using the leave-one-out audit rate strips away the direct effects of a hospital's own behavior, it still includes other hospitals surrounding a given hospital, whose audit rates may still reflect that hospital's behavior. This can be the case if, for example, a given hospital has a large market share. To address this, I consider using the audit rate of other hospitals in the same state in *other* markets, which I define using hospital referral regions. This instrument leverages hospitals whose behavior should not be affected by a given hospital's behavior since they are much farther away in different markets. Similarly, one might be concerned that a hospital's audit rate is correlated with the behavior and audit rates of other hospitals in the same hospital system, as they share a common owner. Figure E16d uses the audit rate of hospitals in the same state but different hospital systems in 2010. The results are robust to using these hospitals to instrument for a hospital's audit rate. Finally, to confirm that the results are not driven by a single state or hospital comparison group, Figure E17 plots the distribution of coefficients when one state or one hospital comparison group is removed from the sample. The coefficients are always negative and the distribution is centered around the main effect.

Finally, I consider a falsification test using state borders in the *interior* of each RAC region. In the interior of each region, there is no change in RAC identity at state borders, so comparing hospitals across these interior borders does not capture exogenous variation driven by different audit strategies across RACs. Figure E18a illustrates the interior borders and the sample of hospitals within one hundred miles of the interior border (excluding hospitals that are within one hundred miles of the RAC border). The falsification test shows no effect on admissions on the "high-audit side" of the interior border (Figure E18b), in contrast to the main results, which show a drop in admissions on the high-audit side of the RAC border.

Patient-Level Analysis In Table F9, I show that the Two Midnight rule difference-in-difference results are robust to varying the sample to include patients who arrive between one and five hours of midnight. Table F5 shows that, in addition to a null effect on revisits within thirty days, there is no effect on revisits within sixty or ninety days.

In column 5 of Table 4, I consider whether there is an effect on non-Medicare patients, who are not directly affected by the Two Midnights rule. I find that after-midnight, non-Medicare ED arrivals do not face a reduction in admissions after the rule is implemented. This indicates that there were no spillovers from the Two Midnights rule onto populations not covered by the rule.

C Rural Hospital Closures

The main results show that RAC audits decrease hospital revenue and increase their costs. This raises the concern that RAC auditing may have driven hospitals into financial distress and, given the prevalence of hospital closures in recent years, led them to close. Hospital closures are associated with decreases in access to care and increases in patient mortality (Carroll, 2019; Gujral and Basu, 2019). To study whether RAC auditing led to hospital closures, I use data from the Sheps Center for Health Services Research on rural hospital closures between 2005 and 2022.[28] I adapt the main specification for the hospital-level analysis to study rural hospital closures. In the border hospital sample, no hospitals closed before 2012 – this is by definition, since the hospital had to be open in 2011 to be audited. Therefore there is no variation in the pre-2010 period to use a difference-in-differences framework. Instead, I run the following specification separately for each year Y in the post period:

$$Close_h^Y = X_h^{2011}\beta^Y + \phi_{g(h)} + \varepsilon_h \tag{10}$$

which regresses a dummy for whether a rural hospital has closed in year Y, $Close_h^Y$, on its (instrumented) audit rate X_h^{2011}, after taking into account the hospital's neighbor comparison group. Figure E13 plots the β^Y coefficients for years where there is variation in closures among rural hospitals in the border sample (i.e., excluding 2012, 2017, and 2021). The results indicate that higher RAC auditing did not cause rural hospitals to close.

D Marginal Value of Public Funds

D.1 Calculations

I next lay out the estimates required to calculate the MVPF in each year. Let θ_t be the estimates on log inpatient revenue in Table 3. Let I_{2010} be a hospital's Medicare inpatient revenue in 2010 (Table 5). Define ΔI_T as the value of the change in cumulative inpatient revenue between 2010 and year T due to an exogenous increase in the audit rate in 2011. If θ_t is the estimated percent reduction in inpatient revenue in year t relative to 2010 (that is, Table 3, column 2), then[29]

$$\Delta I_T = \sum_{t=2011}^{T} \theta_t I_{2010}. \tag{11}$$

The total effect on hospital revenue also includes the money reclaimed back from audits. Let λ_t be the coefficient on payments reclaimed back from hospitals in Figure E11. The eventual value of the reclaimed payments also has to be scaled by the share s of reclaimed payments that were refunded to hospitals in later settlements with hospitals, as discussed in Appendix Section A.2. The value of all the revenue (from deterred admissions and reclaimed payments) returned to Medicare as a result of increasing the 2011 audit rate is:

[28] Data available at https://www.shepscenter.unc.edu/programs-projects/rural-health/rural-hospital-closures/. Last accessed March 2022.

[29] ΔI_T is a negative number because θ_t is negative, and the effect of increased auditing on hospital inpatient revenue is negative.

$$\Delta \text{ hosp. revenue}_T = -\Delta I_T + (1-s)\sum_{t=2011}^{T} \lambda_t. \tag{12}$$

For provider compliance costs up to year T, let K_{2010} be a hospital's 2010 administrative costs (Table 5) and γ_t be the estimated percent increase in compliance costs in year t relative to 2010 (that is, Table 3, column 5). Then

$$\Delta \text{ hosp. compliance costs}_T = \sum_{t=2011}^{T} \gamma_t K_{2010}. \tag{13}$$

The effect on government monitoring costs by year T is defined as the contingency fee f multiplied by the payments reclaimed back from audits (λ_t) in each year between 2010 and T. I assume f to be the midway point between 9 and 12.5: 10.75 percent. If RACs are perfectly competitive and make zero profit, then multiplying by f gives the direct *social* cost of monitoring; otherwise it is an upper bound on the social cost.

$$\Delta \text{ monitoring costs}_T = \sum_{t=2011}^{T} \lambda_t f. \tag{14}$$

The changes in patient health and treatment cost are assumed at baseline to be 0.

$$\Delta \text{ pt health}_T = 0. \tag{15}$$

$$\Delta \text{ treatment costs}_T = 0. \tag{16}$$

The numerator of the MVPF in year T is equal to the sum of the changes in hospital revenue, compliance costs, treatment costs, and patient health, discounted by δ so that it is in terms of 2010 dollars. The denominator is equal to the sum of the changes in hospital revenue (negated) and government monitoring costs, and is also discounted.

D.2 Alternative Assumptions

Figure 8b plots the MVPF in 2013 with alternative assumptions.

Compliance Costs The baseline MVPF calculation uses estimates on compliance costs per year, as measured by hospital administrative costs in HCRIS. I also calculate the MVPF assuming no hospital compliance costs ("No compliance costs"). By 2013 the MVPF would be 1.07 if hospitals did not incur compliance costs, compared to a baseline of 1.42. The lower MVPF reflects that, absent hospital compliance costs, hospitals' willingness to pay to avoid an increase in audit rate would be lower.

Medicare Savings The baseline calculation assumes that Medicare saves all the revenue from deterred admissions. Figure 8b shows the results for alternative assumptions about Medicare savings. If there is no deterrence effect and all the savings are from the reclaiming of denied payments ("Savings: denials only"), then the MVPF of RAC audits is very high (4.55 by 2013), making it a much less attractive source of government revenue. I also consider alternate assumptions that Medicare only saves *some* of the revenue from deterred admissions. First, I consider making the assumption that all deterred inpatient stays become observation stays ("Savings: inpt $\rightarrow$ obs"). I assume that each hospitals would be paid \$3,160 for each observation stay that substitutes for an inpatient stay, which I take from a MedPAC report on the difference in Medicare payments for inpatient stays and comparable observation stays (Medicare Payment Advisory Commission (2015), Figure 7-2). This results in a slightly higher MVPF than baseline (1.66 by 2013), as the government's savings are smaller since it has to pay for observation stays. Second, I consider using the estimates on the effect of total inpatient *and* outpatient revenue from Figure E12 ("Savings: inpt + outpt"). This leads to a slightly lower MVPF than baseline (1.29 by 2013), as the estimates on savings from combined inpatient and outpatient revenue are more pronounced than just the savings from inpatient revenue alone.

Patient Health Effects The baseline calculation assumes that there are no patient health effects for the marginal patient denied a hospital admission. I relax this assumption by considering two possible scenarios: first that the marginal admission increases patient mortality ("Patient health $\uparrow$"), and second that it decreases patient mortality ("Patient health $\downarrow$"). I take the estimates of the effect of the marginal hospital admission on mortality from Currie and Slusky (2020), which reports a (statistically insignificant) 0.457pp *increase* in 7-day mortality and a 0.488pp *decrease* in 15-day mortality for the marginal hospital admission. Using a value of a statistical life of \$1 million, I find that the MVPF calculations are sensitive to the assumption made about effects on patient health. However, given that neither my nor Currie and Slusky's (2020) analysis finds a statistically significant effect on patient health, so I assume at baseline no effect on patient health effects.

Treatment Cost The baseline MVPF calculation assumes, conservatively, that there is no change in treatment cost when inpatient stays are deterred by RAC audits. I relax this assumption two ways: first, by assuming that Medicare payments for inpatient stays are a constant markup of the costs ("Treatment costs: "markup") , and second, by assuming that inpatient stays become observation stays, and Medicare payments for observation stays are a markup of the treatment costs ("Treatment costs: obs"). The first scenario assumes that Medicare inpatient payments are a 1.55x markup of hospital costs, which is based on estimates of markups for one-day stays by MedPAC (2015). The second scenario assumes that Medicare payments for observation stays are 30 percent that of payments for inpatient stays,[30] and the Medicare payments for observation stays are a 1.55x markup of the hospital's costs to provide an observation stay. Both scenarios are less conservative than the baseline assumption of no changes in treatment costs and result in lower MVPFs than baseline.

E Extrapolation to Overall Hospital Sample

This section describes the calculation to extrapolate the savings estimates from the border hospital sample to the overall RAC program. This calculation rests on fairly strong assumptions, but nonetheless may be of interest for gauging the magnitude of overall savings from the RAC program. First, we must assume that the savings scale linearly with audit rate, so that the effects estimated from a marginal increase in audit rate can be extrapolated beyond the support to a wide range of audit rates. Second, we must assume homogeneous treatment effects across hospitals in the border sample and overall hospitals. Note that while hospitals on opposite sides of the border are similar to each other (Table F2), the border hospital sample differs from the overall sample. Hospitals in the border sample are smaller, more rural, and disproportionately from the Midwest RAC region, Region B, (Table F1). Third, this calculation assumes that even at high levels of auditing, there is still no effect on other outcomes that may affect welfare, like patient health or hospital closures.

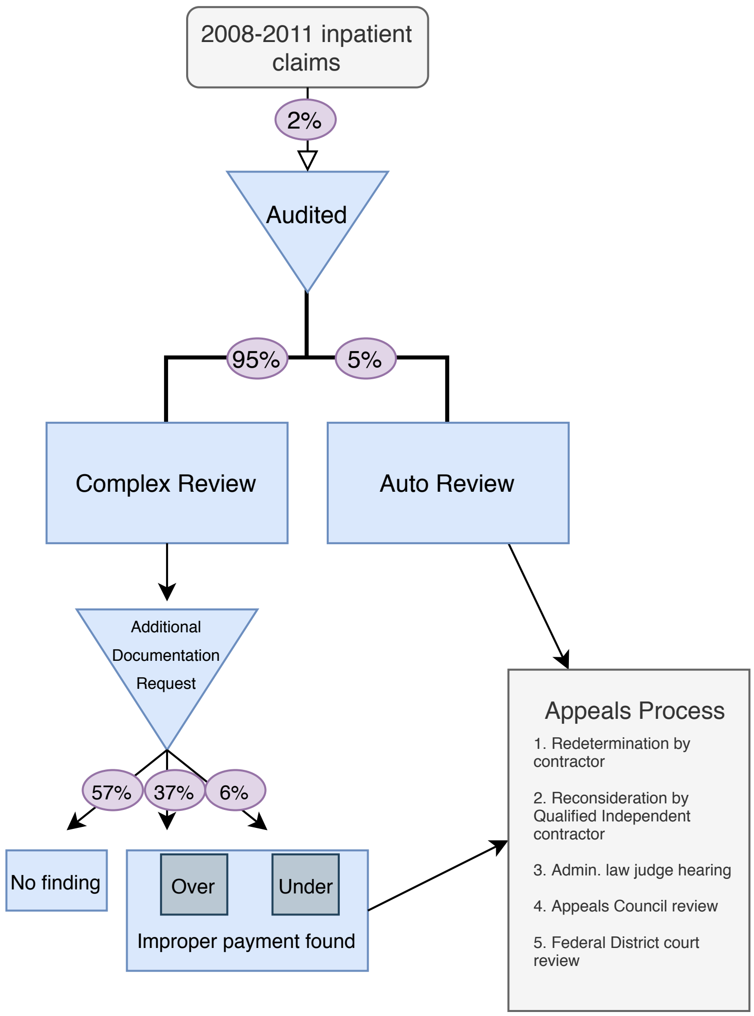
Under these assumptions, I can calculate the extrapolated savings by multiplying the 2011-2015 event study coefficients on Medicare inpatient revenue (Figure 4b) and payments demanded (Figure E11) by each hospital's 2011 audit rate. Since the estimates are based on the logarithm of inpatient revenue and represent a percent change relative to the baseline in 2010, I multiple these coefficients by the hospital's 2010 inpatient revenue. Figure E21 plots the extrapolated savings for each hospital-year, compared to the actual changes in Medicare inpatient revenue and actual payments demanded. For both types of savings, the extrapolated and actual savings are positively correlated. This indicates that in the overall sample, hospitals subject to higher audit rates reduced their Medicare inpatient revenue and were subject to more audit demands in subsequent years. Summing up the extrapolated savings across all hospitals from 2011 to 2015 implies that the RAC program saved the Medicare program \$9.28 billion between 2011 and 2015, compared to the actual \$11.74 billion in savings from reductions in inpatient spending and audit demands in this period. Note, however, the relatively low R^2 from the regression between extrapolated and actual savings, indicating that much of the variation in savings is not explained by variation in 2011

[30] The ratio between the average 2010 payment for an inpatient stay (\$5640) for hospitals in the border sample and the average payment for an observation stay in the same sample (\$1671) is 0.296.

audit rate.

F Appendix Figures

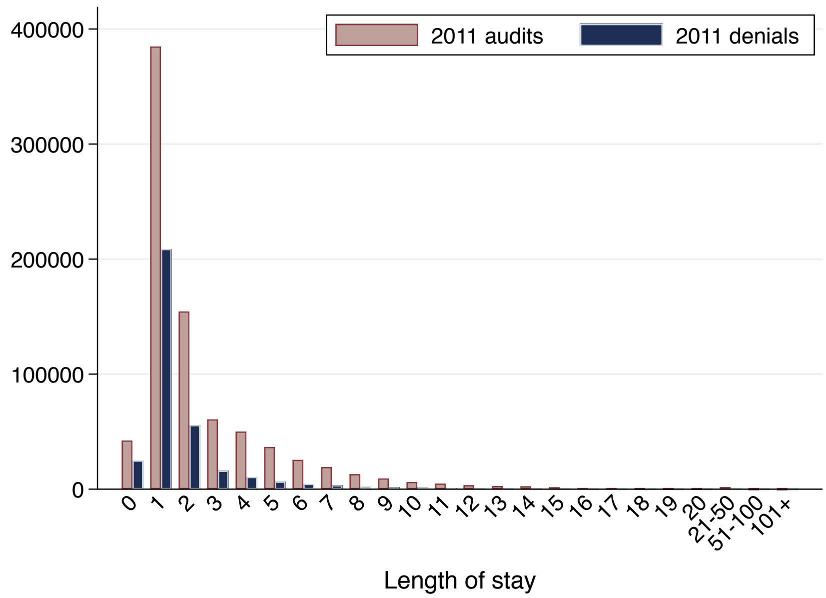
Figure E1. RAC Inpatient Claims Auditing and Appeals Process, 2011 Audits



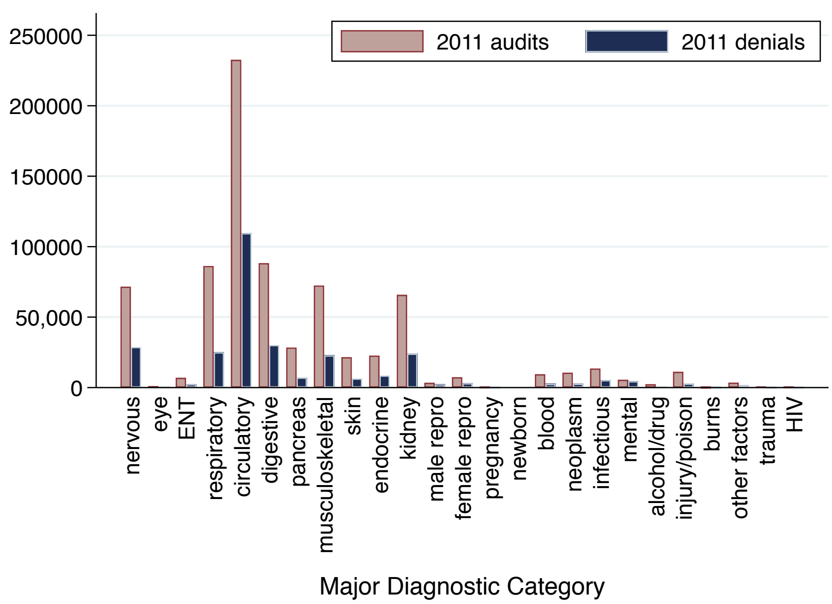
This figure illustrates the stages of the claims auditing and appeals process. The percentages in ovals denote the percent of claims that, conditional on reaching a given stage in the process, reach the next stage. The percentages are calculated based on audits in 2011 of inpatient claims between 2008 and 2011. Data: CMS audit data.

Figure E2. 2011 Audit and Denial Counts by Stay Characteristics

(a) By Length of Stay

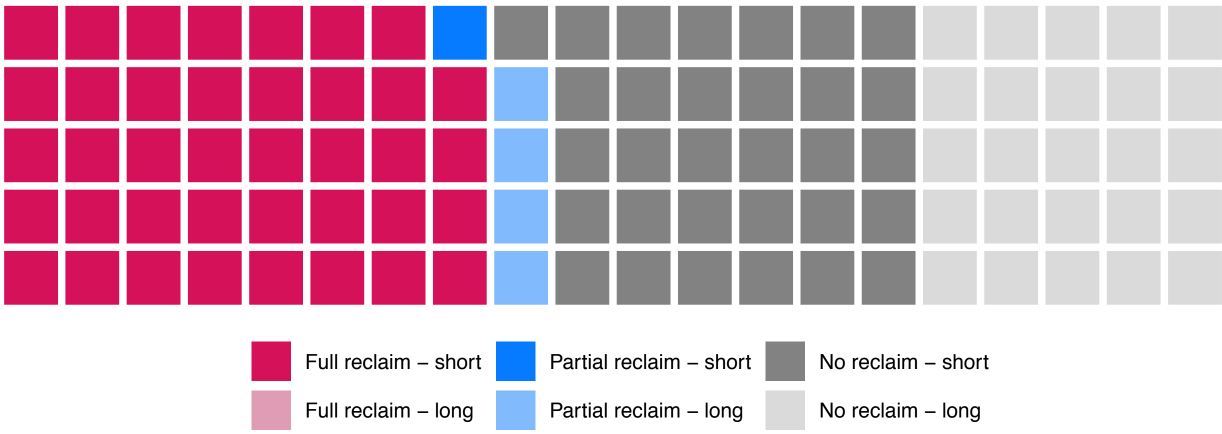


(b) By Major Diagnostic Category



This figure plots the count of 2011 audits and denials by (a) an admission's length of stay and (b) the Major Diagnostic Category associated with the admission's DRG. Data: MEDPAR and CMS audit data.

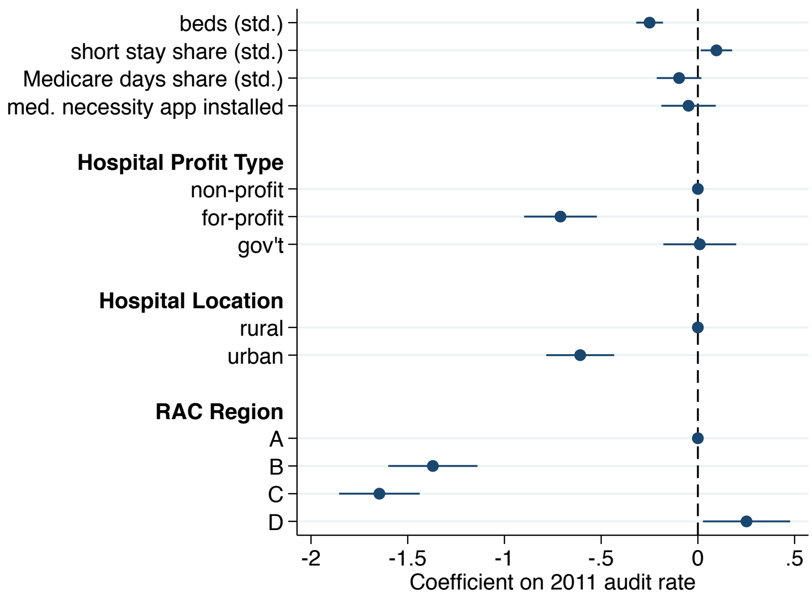
Figure E3. 2011 Audit and Denial Characteristics



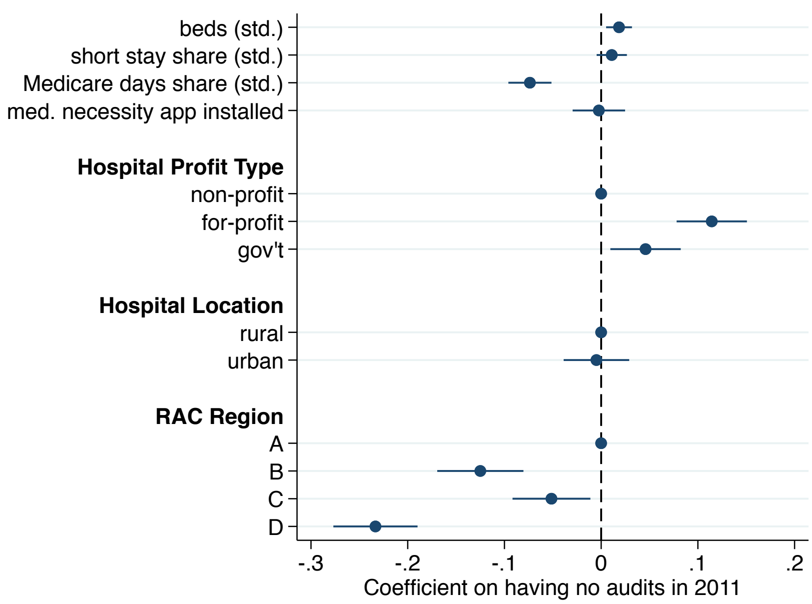
This figure is a waffle plot of 2011 audits of inpatient stays in 2008-2011, where each box represents one percent of total audits. The dark shaded boxes of each color denote audits of inpatient stays. The red and blue colored boxes denote audits that result in the full payment being reclaimed or a partial payment being reclaimed, respectively. The figure plots the following shares of 2011 inpatient stay audits: 39 percent of audits are for short stays where the full payment is reclaimed, less than 1 percent of audits are for long stays where the full payment is reclaimed, one percent of audits are for short stays where a partial payment is reclaimed, 4 percent of audits are for long stays where a partial payment is reclaimed, 31 percent of audits are for short stays where there is no payment reclaimed, and 25 percent of audits are for long stays where there is no payment reclaimed. Data: MEDPAR and CMS audit data.

Figure E4. Correlation between Hospital Characteristics on 2011 Audit Rate and No Audit

(a) Outcome: 2011 hospital audit rate

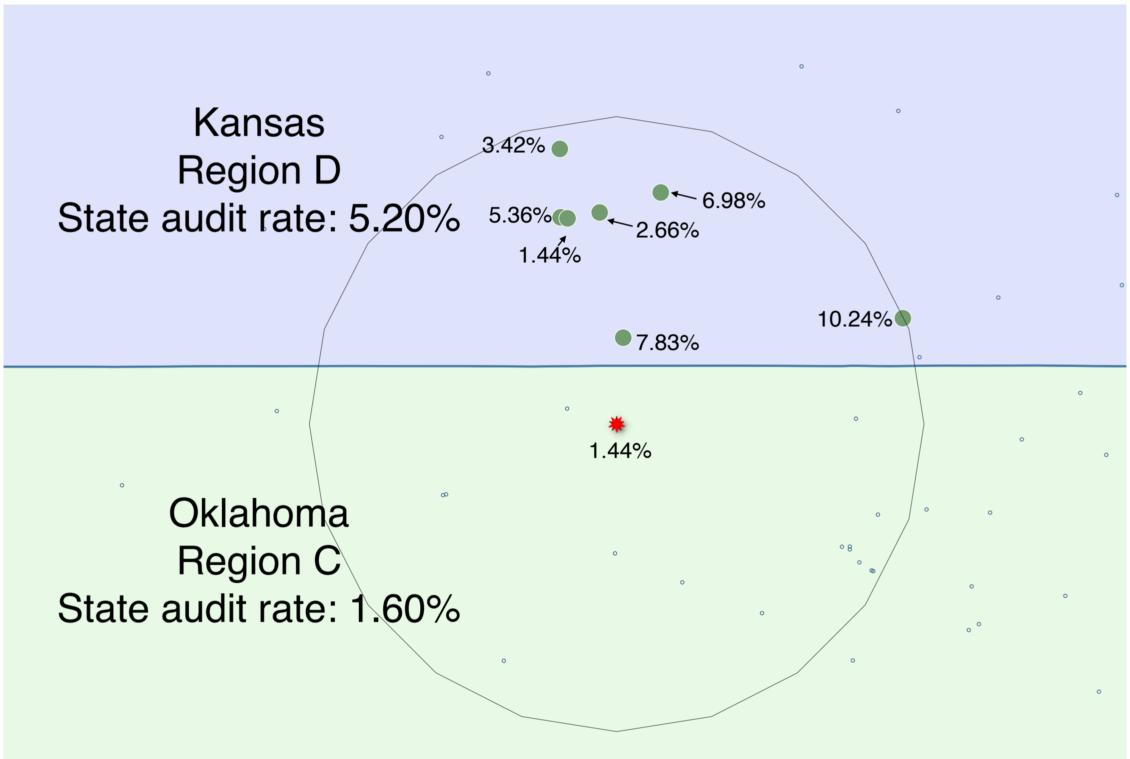


(b) Outcome: no audits at hospital in 2011



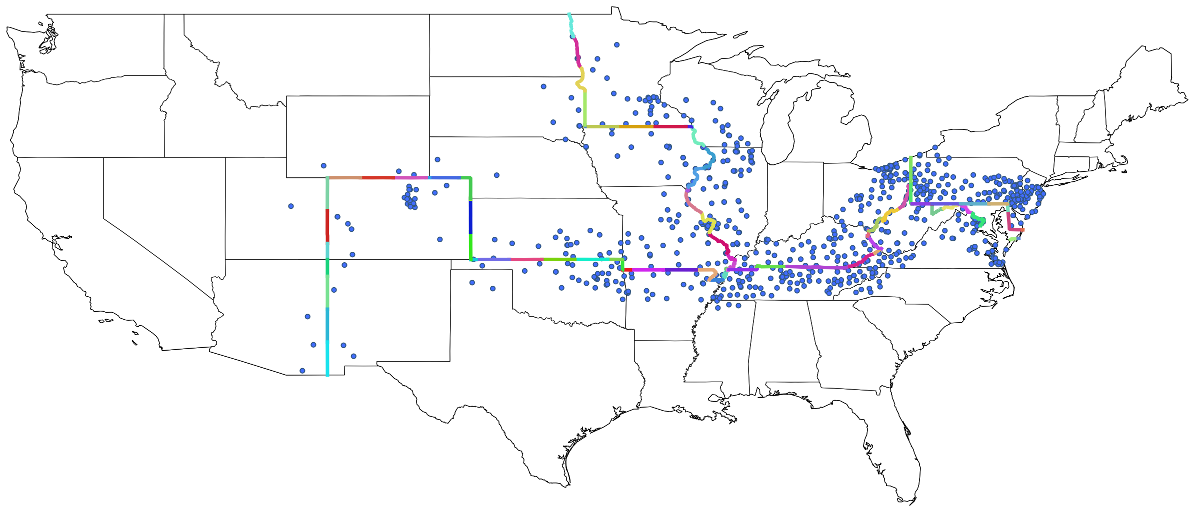
These figures plot coefficients from a regression of (a) a hospitals 2011 audit rate and (b) an indicator variable for whether a hospital was not audited in 2011 on 2010 hospital characteristics. Short stay share is the share of 2010 Medicare admissions with lengths of stay 0-2. Medicare days share is percent of hospital days that are Medicare. Beds, short stay share, and Medicare days share are standardized relative to the mean. Data: MEDPAR, CMS audit data, and Medicare Provider of Services file.

Figure E5. Example of Border Hospital and Neighbor Comparison Group Definition



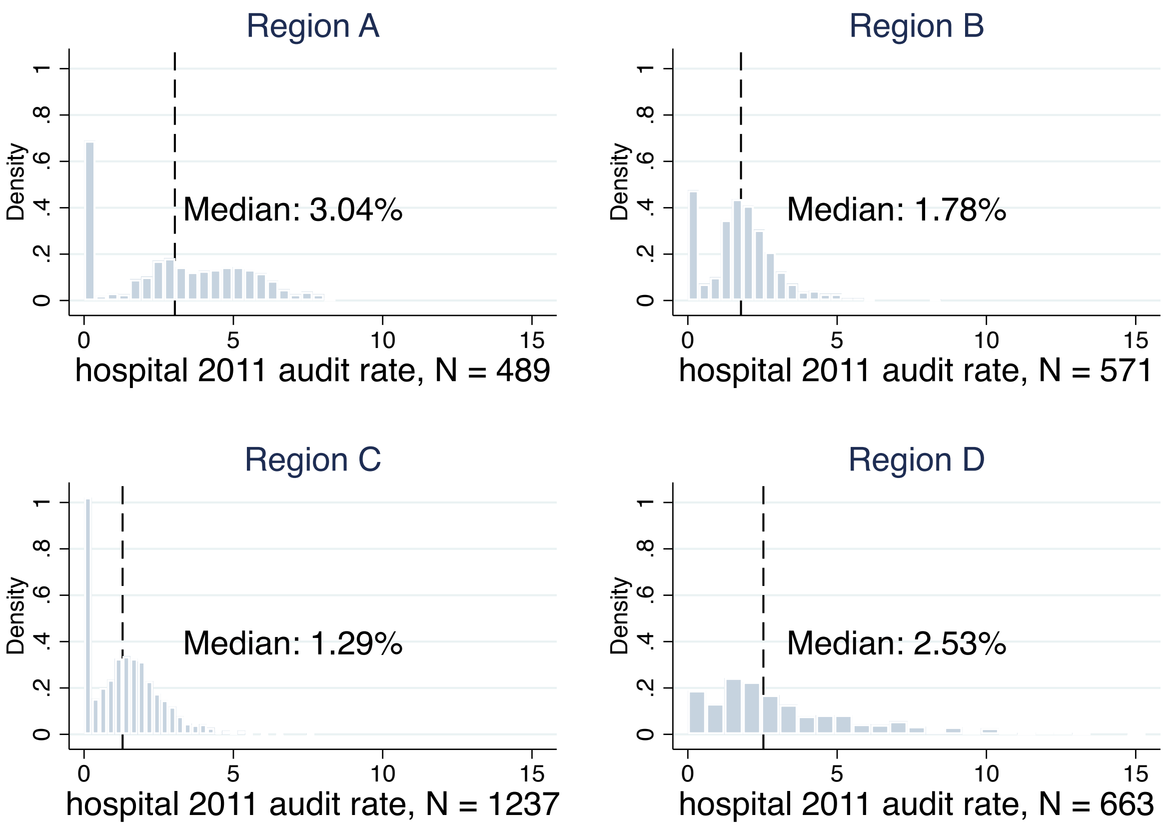
This figure illustrates how a "neighbor comparison group" is identified for each border hospital in the across-hospital empirical strategy. Neighboring hospitals are all hospitals within a 100 mile radius of a hospital, on the opposite side of the RAC border. In this example, the green circle hospitals in Kansas are considered neighboring hospitals to the red spiked hospital in Oklahoma.

Figure E6. RAC Border Segments and Hospitals Within 100 Miles



This figure shows how the RAC border is divided into one-hundred mile segments that do not cross state borders, and all hospitals within one-hundred miles of the RAC border.

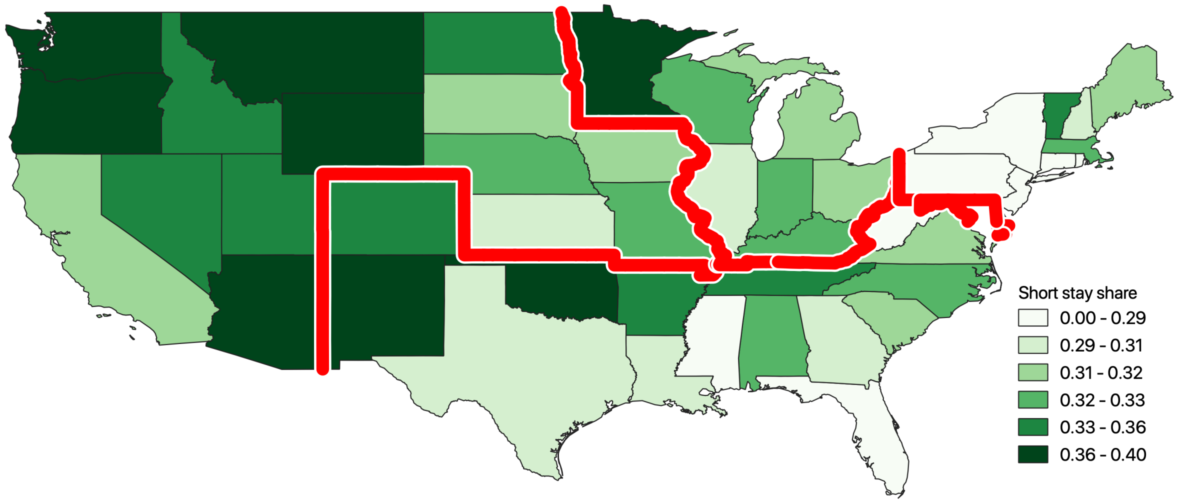
Figure E7. Histogram of 2011 Hospital Audit Rates by RAC Region



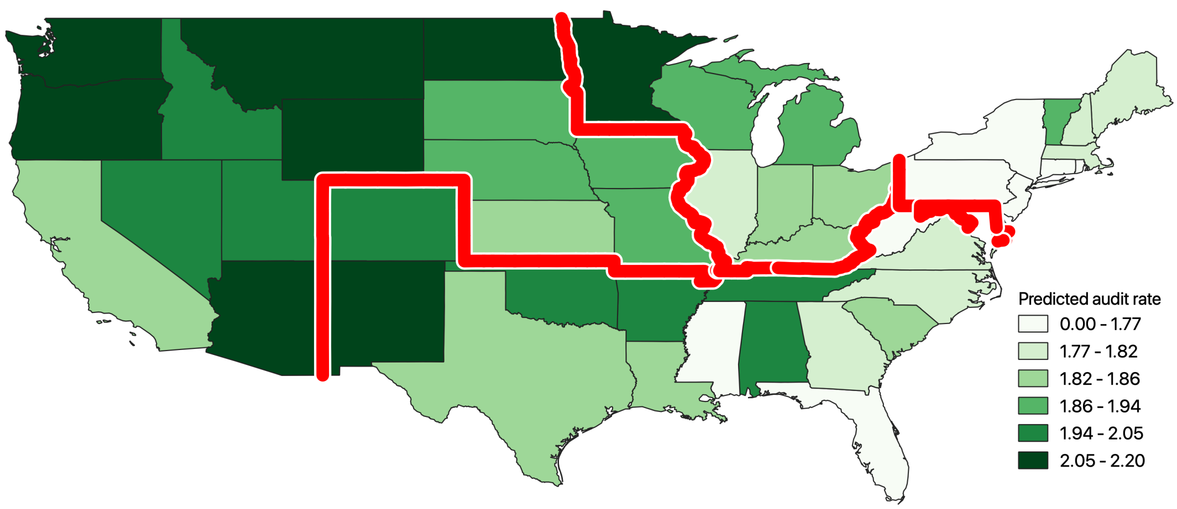
This figure plots the histogram of 2011 hospital audit rates by RAC region, where audit rate is defined as the percent of a hospital's 2008-2011 claims that were audited by RACs. Data: MEDPAR and CMS audit data.

Figure E8. 2010 Average Short Stay Share of Medicare Admissions and Predicted 2011 Audit Rate by HRR

(a) 2010 Average Short Stay Share of Medicare Admissions by State

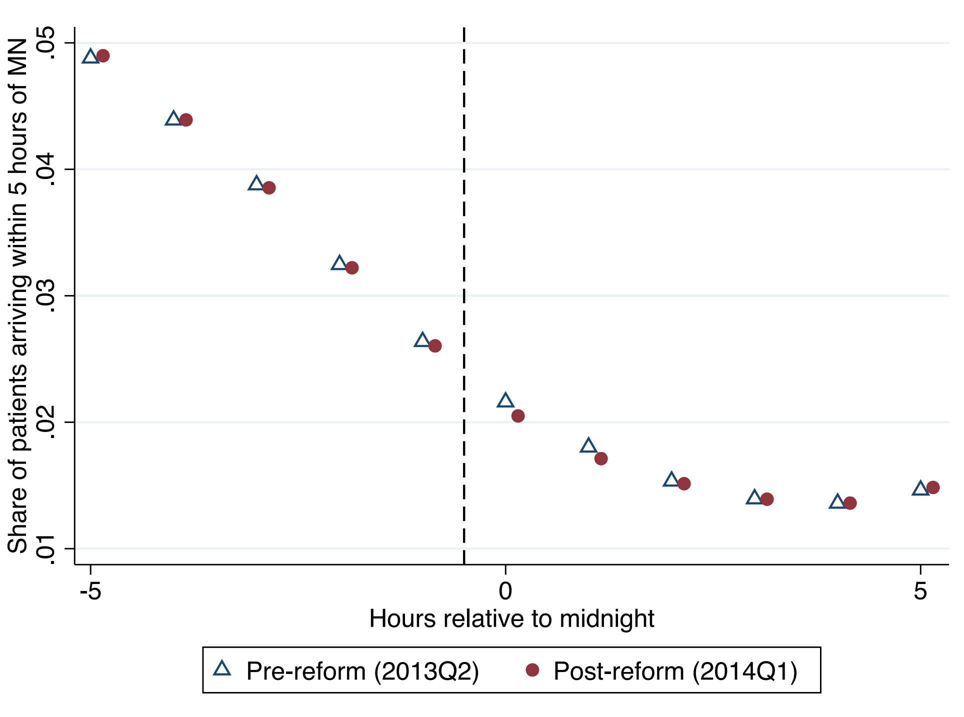


(b) Predicted 2011 Audit Rate by State



These figures plot state averages of hospital-level characteristics. The top panel plots the average share of Medicare admissions with a length of stay of 0-2 in 2010, and a darker shade is associated with a higher share. The bottom panel plots the predicted 2011 audit rate using characteristics of 2007-2009 claims. The prediction specification is a regression of the likelihood of being audited in 2011 on admission month, major diagnostic category, admission source, and length of stay for each hospital's 2007-2009 claims. The red line demarcates RAC regions, which are: Region A (Northeast), Region B (Midwest), Region C (South), and Region D (West). Darker shades denote higher audit rate. The red line demarcates RAC regions. Maryland was not audited under the RAC program as it uses a unique all-payer rate-setting system for hospital services. Data: MEDPAR and CMS audit data.

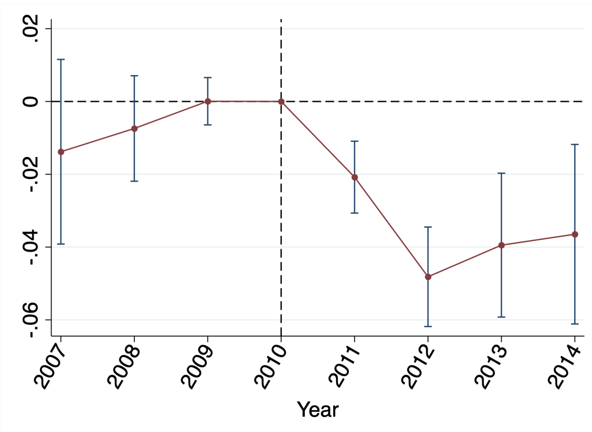
Figure E9. Share of Medicare ED Patients By Hour of ED Arrival



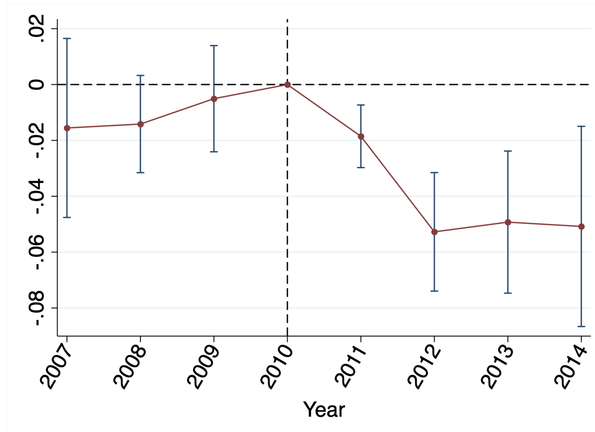
This figure plots the share of Medicare patients that arrive at the ED at each hour (relative to midnight) pre- and post-reform, among traditional Medicare patients who arrived in the ED within 5 hours of midnight in Florida. Data: HCUP SID/SEDD.

Figure E10. Event Studies of Effect of 2011 Audit Rate on Medicare Admissions and Revenue, by Diagnosis

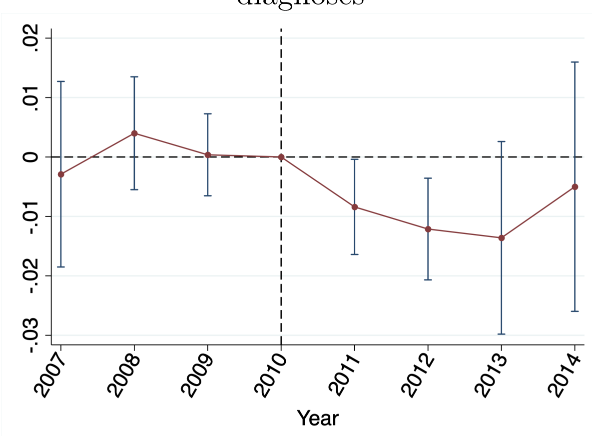
(a) Log Medicare admissions, circulatory diagnoses



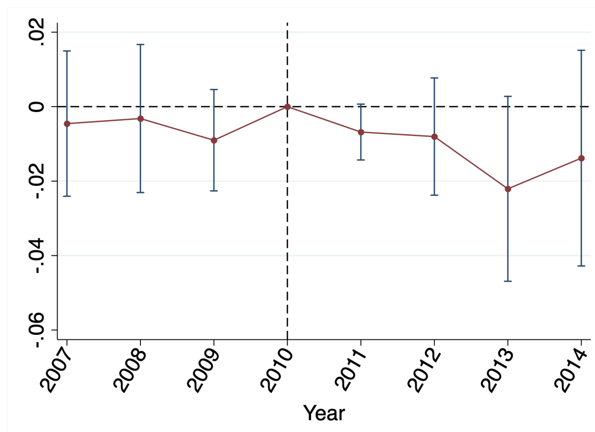
(b) Log Medicare inpatient revenue, circulatory diagnoses



(c)
Log Medicare admissions, non-circulatory diagnoses

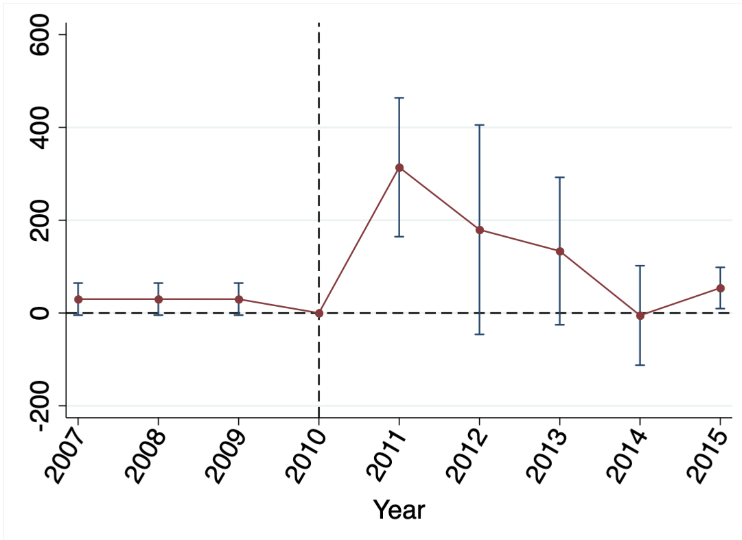


(d) Log Medicare inpatient revenue, non-circulatory diagnoses



This figure plots event studies of the reduced form coefficients and 95% confidence interval in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome. Medicare admissions and revenue are from MEDPAR. Circulatory diagnoses are identified by the Major Diagnostic Category of the claim's diagnosis related group (DRG). Since the mapping from DRGs to MDCs depends on ICD-9 diagnosis codes, which were phased out in 2015, the time period for the event study is restricted to 2007-2015. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

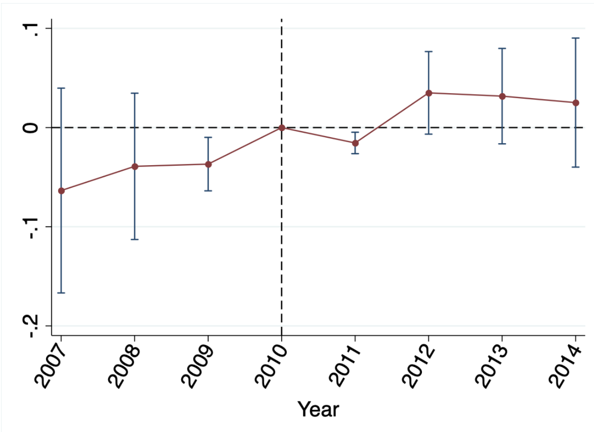
Figure E11. Event Study on Effect of 2011 Audit Rate on Payment Demanded ($1000s) from RAC Audits



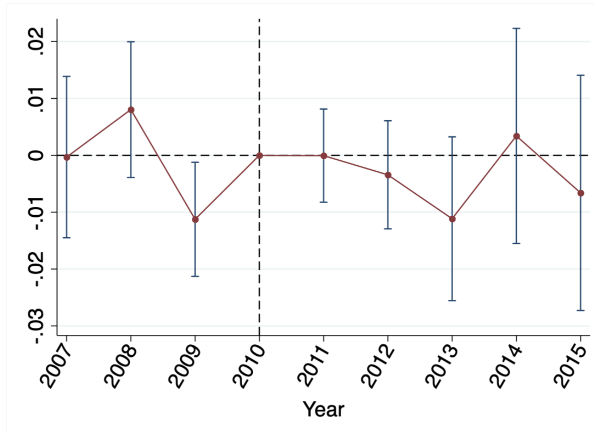
This figure plots event studies of the reduced form coefficients and 95% confidence interval in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome. The outcome is the amount of payment demanded initially from RAC audits of inpatient stays, by year of audit. Data: CMS audit data.

Figure E12. Event Studies on Effect of 2011 Audit Rate on Hospital Outpatient Revenue and Observation Stays

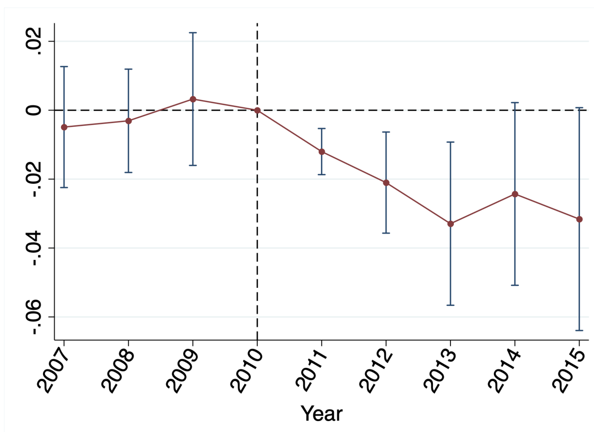
(a) Log observation stays



(b) Log Medicare outpatient revenue

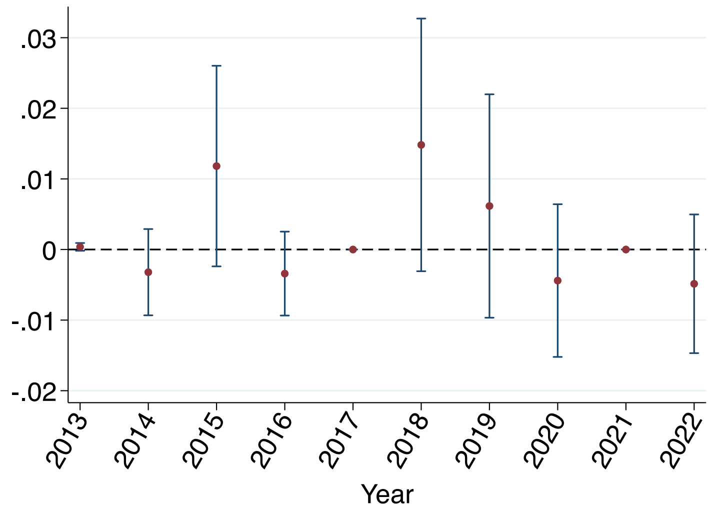


(c) Log Medicare inpatient and outpatient revenue



This figure plots event studies of the reduced form coefficients and 95% confidence interval in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome. Observation stays are defined as outpatient claims associated with revenue center "0760" or "0762," or the HCPCS procedure codes "G0378" or "G0379." Outpatient revenue is the sum of all Medicare outpatient payments. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

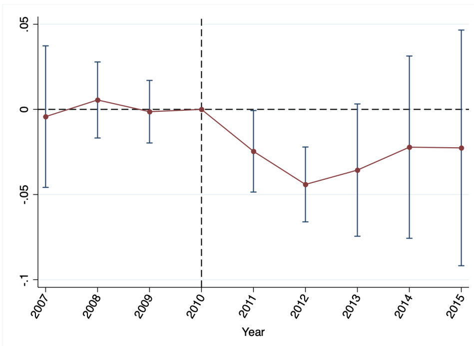
Figure E13. Coefficients of Effect of 2011 Audit Rate on Rural Hospital Closure in a Given Year



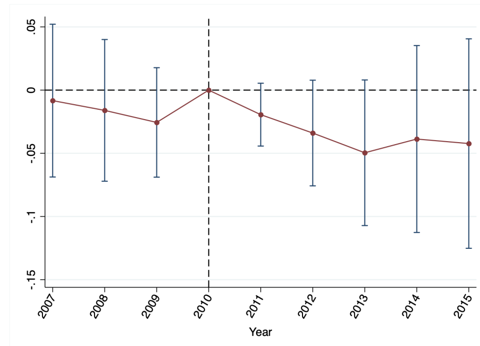
This figure plots the coefficients from individual regressions of the instrumented 2011 audit rate on a dummy for whether a hospital closed in a given year, for rural hospitals in the border sample. There are no closures prior to 2013 and no closures in 2017 and 2021 in the border hospital sample. Data: Sheps Center for Health Services Research and CMS audit data.

Figure E14. Event Studies on Effect of 2011 Denial Rate on Medicare Admissions and Revenue, and Administrative Burden

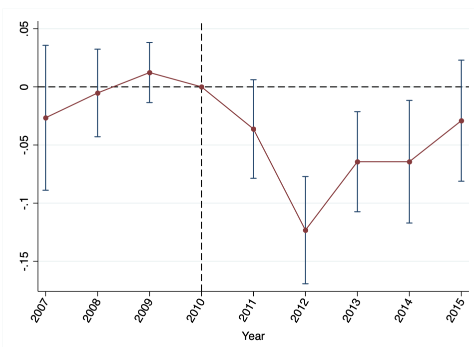
(a) Log Medicare admissions



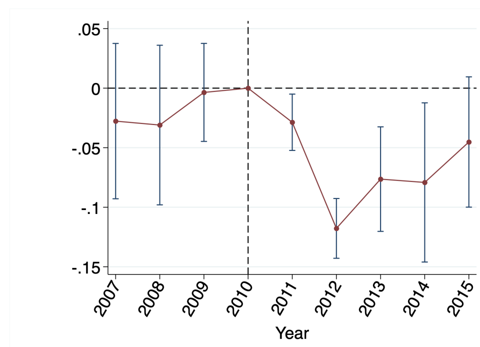
(b) Log Medicare inpatient revenue



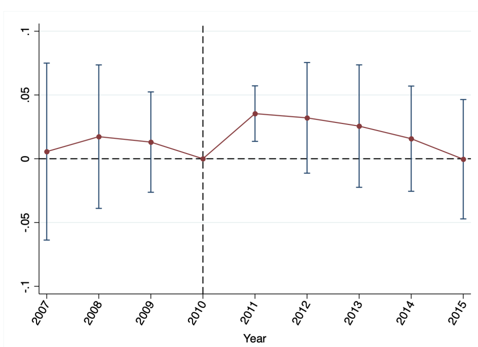
(c) Log Medicare admissions, LOS ≤ 2



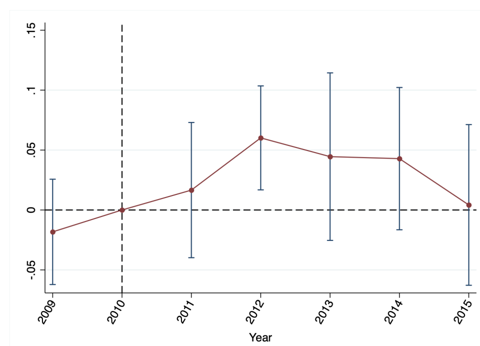
(d) Log Medicare inpatient revenue, LOS ≤ 2



(e) Log net administration costs



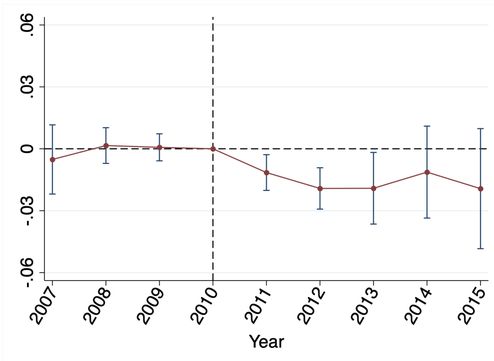
(f) Indicator for installing medical necessity software



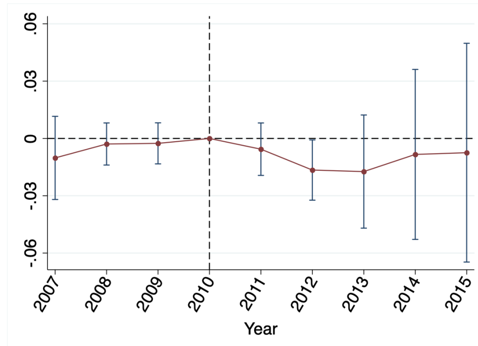
This figure plots event studies of the reduced form coefficients and 95% confidence interval in Equation 3 (using the denial rate rather than the audit rate), scaled by the correlation between the leave-one-out 2011 denial rate and the actual 2011 denial rate in the weighted border hospital sample. Denial rate is the share of claims that are audited and result in an overpayment demand or repayment for an underpayment. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 denial rate rate on a hospital-level outcome. Medicare admissions and revenue are from MEDPAR. Inpatient revenue is the sum of all Medicare inpatient payments. Net administration costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Indicator for installing software is equal to 1 if a hospital reports the status of a medical necessity software as "contracted/not yet installed," "installation in process," and "to be replaced" in HIMSS. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

Figure E15. Robustness to Sample Definition: Event Studies on Effect of 2011 Audit Rate on Log Medicare Admissions

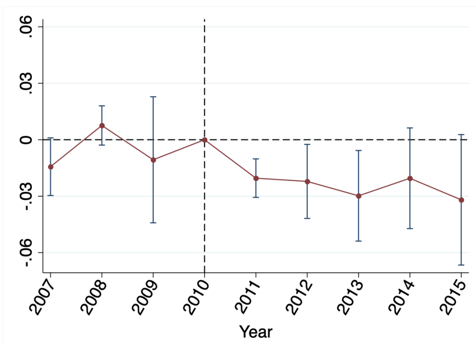
(a) Sample: within 100 miles (main)



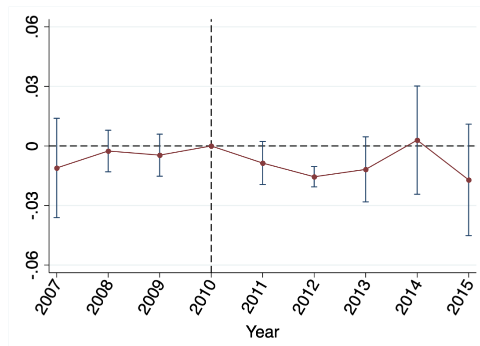
(b) Sample: within 50 miles



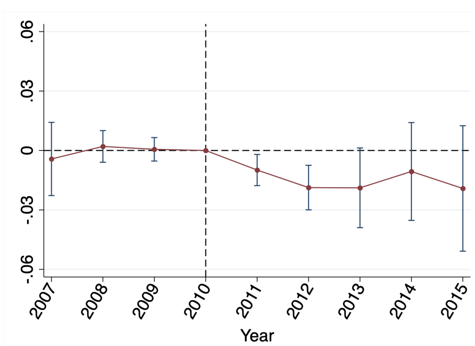
(c) Sample: within 150 miles



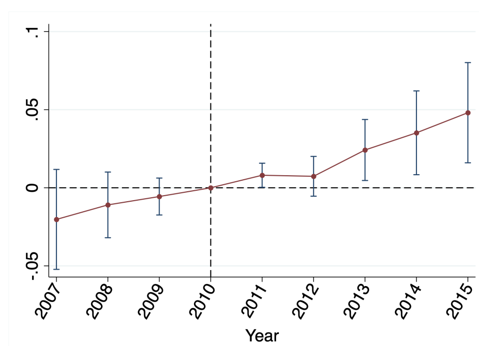
(d) Sample: w/in 100 miles, outside 10



(e) Sample: within 100 miles, audit rate > 0

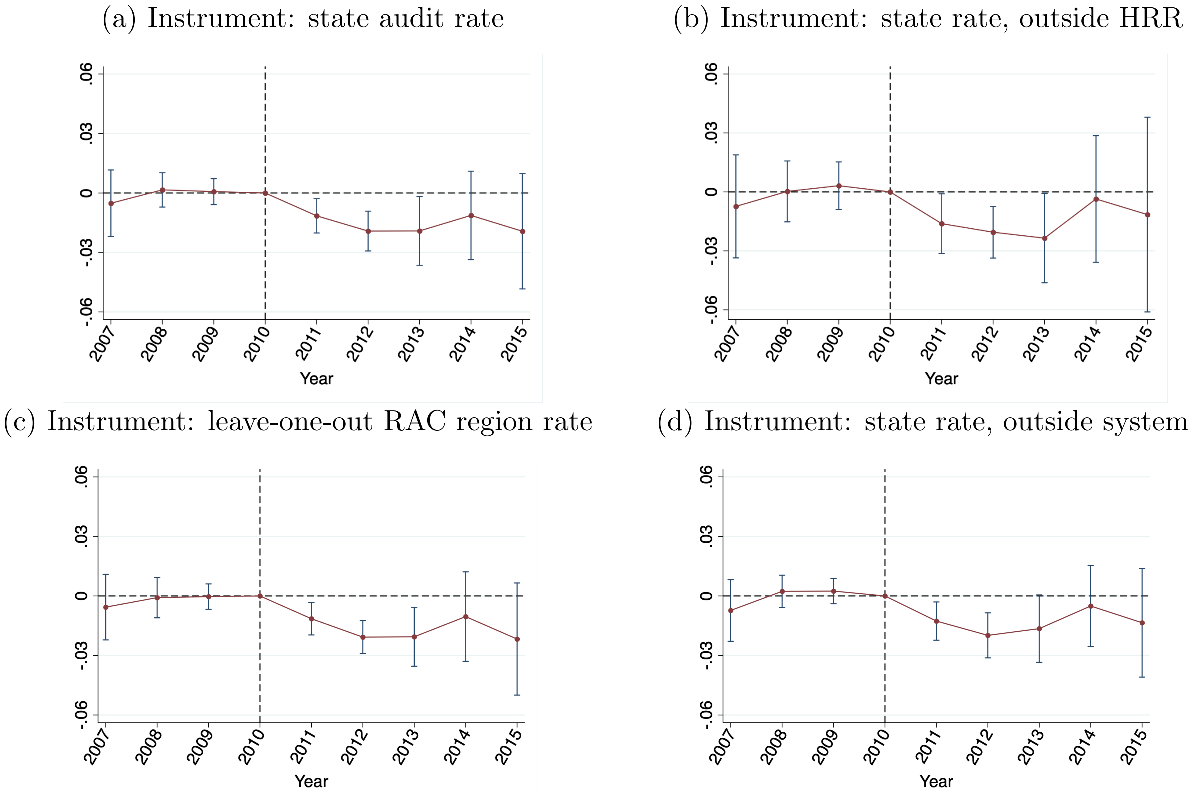


(f) Sample: all hospitals



This figure plots robustness analysis event studies of the scaled reduced form coefficients and 95% confidence intervals of the specification in Equation 3, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient estimates the effect of a one percentage point increase in 2011 audit rate on log Medicare admissions. The figures plot the results using different definitions of the border sample: (a) reproduces the main result and defines the border sample to be all hospitals within 100 miles of the RAC border; (b) defines the border sample to be all hospitals within 50 miles of the RAC border, (c) defines the border sample to be all hospitals within 150 miles of the RAC border, (d) defines the border sample to be all hospitals within 100 miles of the RAC border, excluding hospitals within 10 miles of the border, and (e) uses the 100 mile border sample and restricts to hospitals with 2011 audit rate greater than 0. Panel (f) plots the results for all hospitals (N=3014), in a specification where the hospitals audit rate is instrumented using the leave-one-out RAC region rate and includes hospital and year fixed effects. Data: MEDPAR.

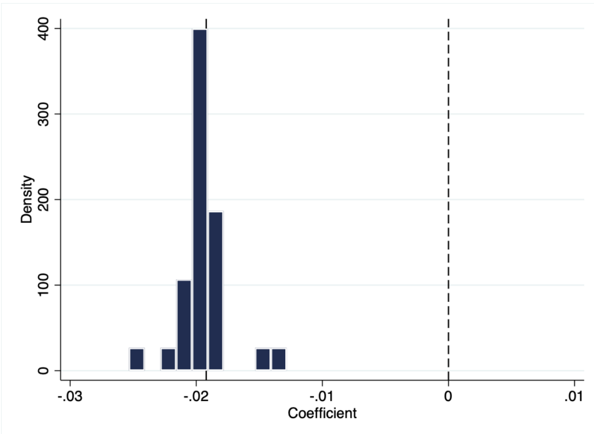
Figure E16. Robustness to Instrument Definition: Event Studies on Effect of 2011 Audit Rate on Log Medicare Admissions



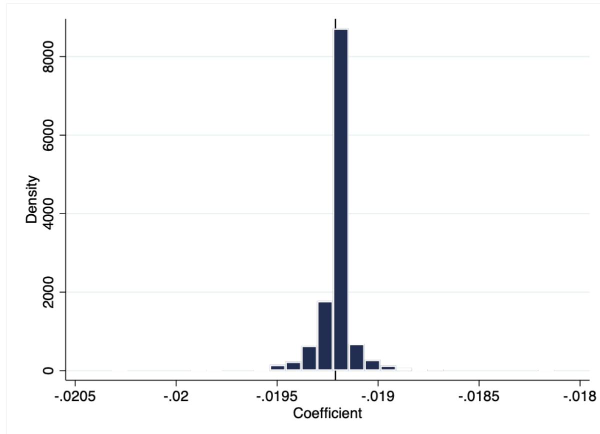
This figure plots robustness analysis event studies of the reduced form coefficients and 95% confidence intervals of the specification in Equation 3, scaled by the correlations between the instruments and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on log Medicare admissions. The figures plot the results using different instruments for a hospital's 2011 audit rate. Panel (a) uses 2011 state audit rate and panel, (b) uses 2011 audit rate among hospitals in the same state but in different hospital referral regions (HRR) as the hospital, (c) uses the 2011 audit rate of other hospitals in the same RAC region, and (d) uses the 2010 audit rate of other hospitals in different hospital systems in 2010. Data: MEDPAR and hospital systems from Cooper et al. (2019).

Figure E17. Robustness Test: Leave-One-Out Coefficients of 2012 Effect of 2011 Audit Rate on Log Medicare Admissions

(a) Leaving 1 state out at a time



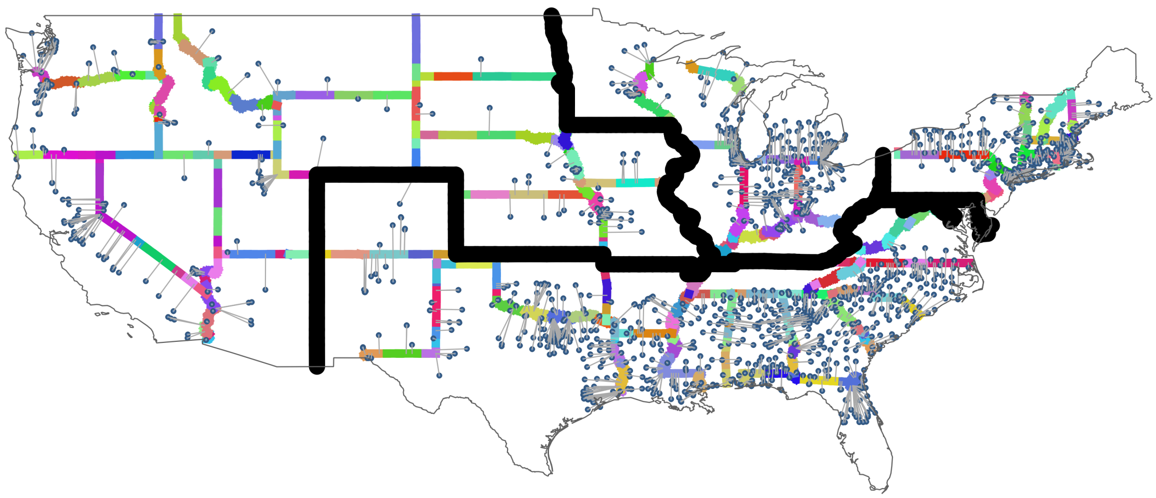
(b) Leaving 1 comparison group out at a time



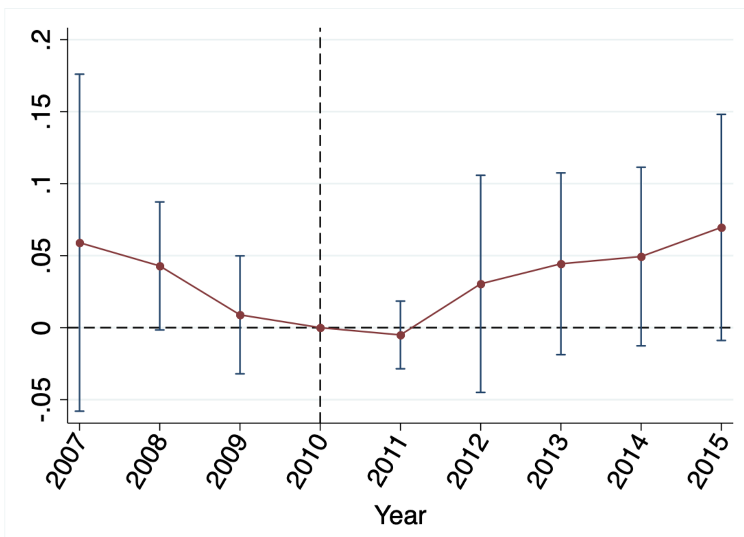
This figure plots distributions of the 2012 coefficient of the reduced form event study specification in Equation 3 on log Medicare admissions, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample the outcome. Panel (a) plots the distribution of the coefficient when leaving one state out at a time, and panel (b) plots the distribution of the coefficient when leaving one hospital neighbor comparison group out at a time.

Figure E18. Falsification Test: Interior State Borders

(a) Falsification Test Border Segments and Hospitals Within 100 Miles



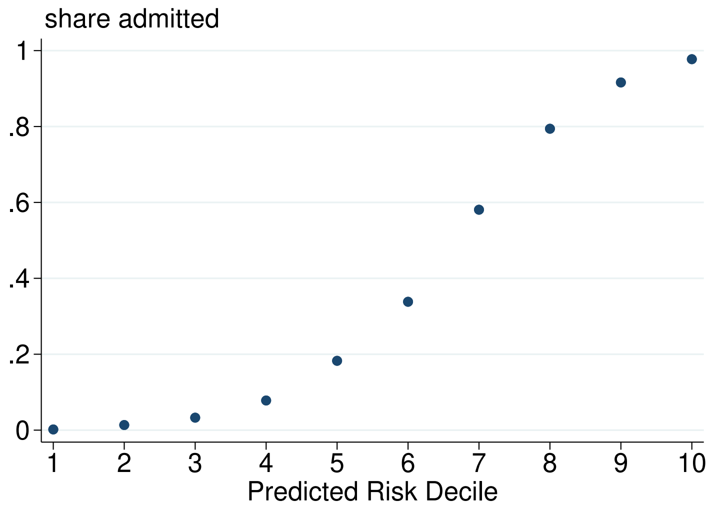
(b) Event Study on Effect of 2011 Audit Rate on Log Medicare Admissions



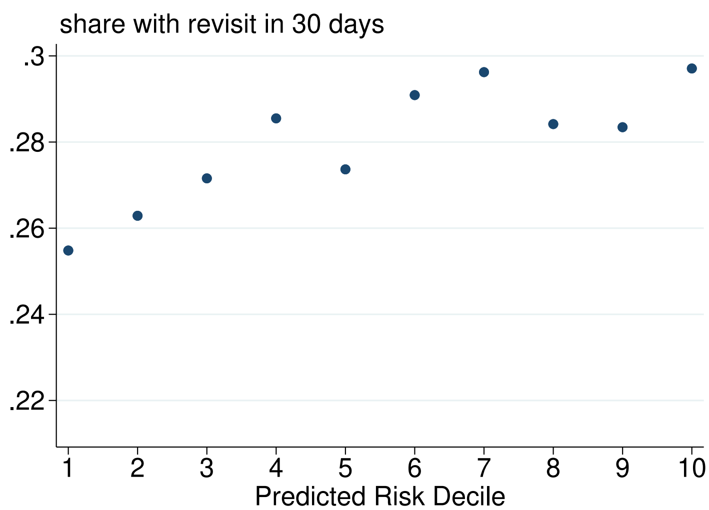
The top panel of this figure plots a map of state borders on the interior of RAC regions, divided into 100-mile segments that do not cross state borders. The RAC border is the thick black line. Each dot represents a hospital within 100 miles of the interior state borders, excluding hospitals that are in the main sample (within 100 miles of the RAC border). The line between the hospital and the interior state border denotes the closest interior state border to that hospital. The bottom panel plots the reduced form coefficient and 95% confidence interval of the specification in Equation 3 (scaled by correlation between 2011 audit rate and 2011 leave-one-out audit rate in the interior border hospital sample), where the outcome variable is log Medicare admissions (MEDPAR). Sample is comprised of hospitals within 100 miles of the state interior border with at least 1 hospital in their "neighbor hospital comparison group" and are clustered at the state and border segment level.

Figure E19. Average Outcomes by Patient Severity

(a) Inpatient

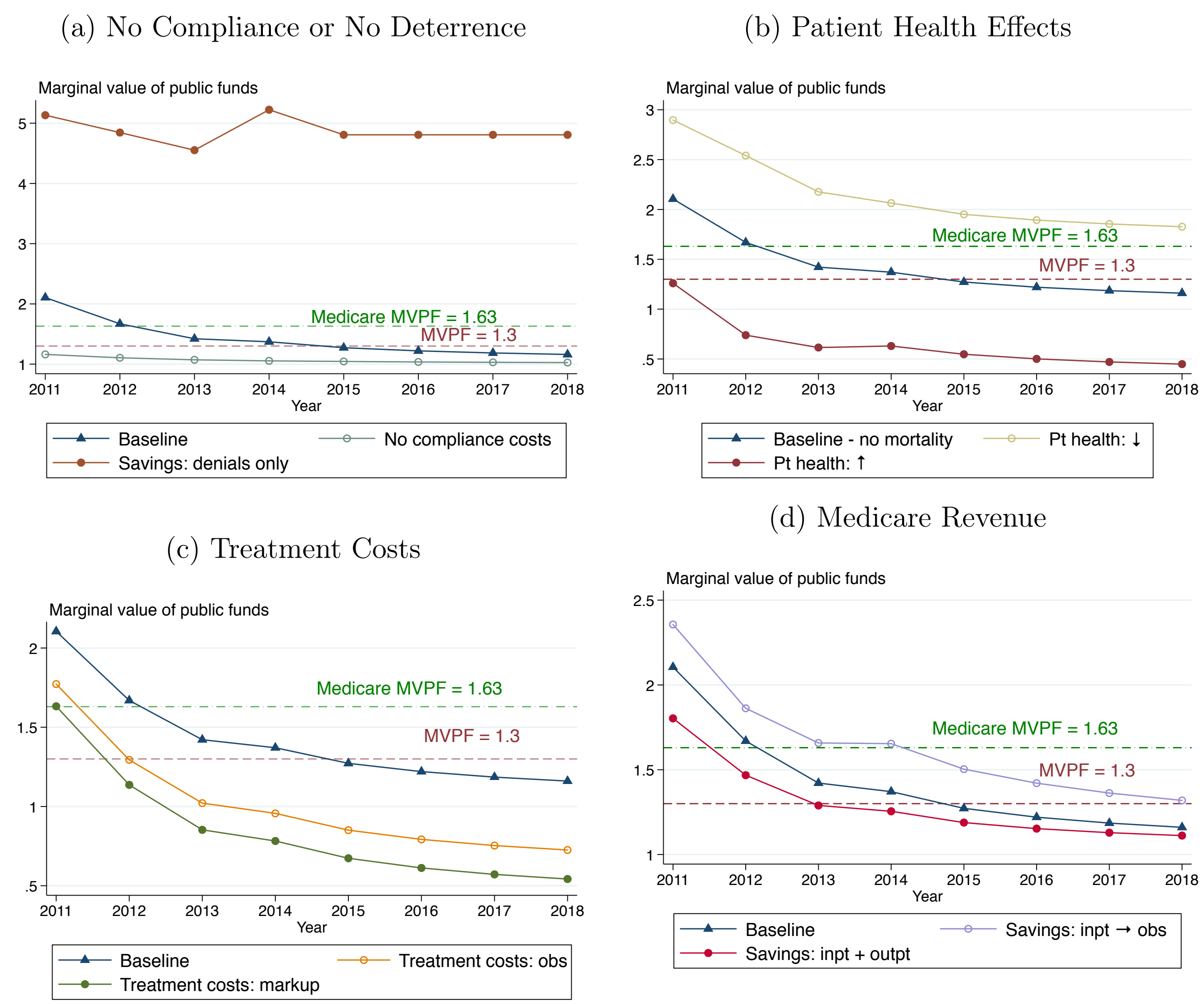


(b) Revisit in 30 days



This figure plots (a) the share of patients admitted as inpatient from the ED and (b) the share of patients with a revisit within 30 days by predicted severity decile, in 2013Q2. Patient risk is predicted by estimating a logit using ED visits between 9:00AM and 3:00PM of an indicator for being admitted within 30 days of an ED visit on patient demographics, current ED visit information, and information on any prior visits in the last 365 days. Demographics include age-bin, sex, race, Hispanic indicator, point of origin indicator, and mean zip code income. Information on current visit includes hospital and quarter. Information on previous visits includes the number of visits/inpatient stays/length of stay in the last month or last year, as well as any diagnoses and procedures recorded in stays within the last month or last year. Data: HCUP SID/SEDD.

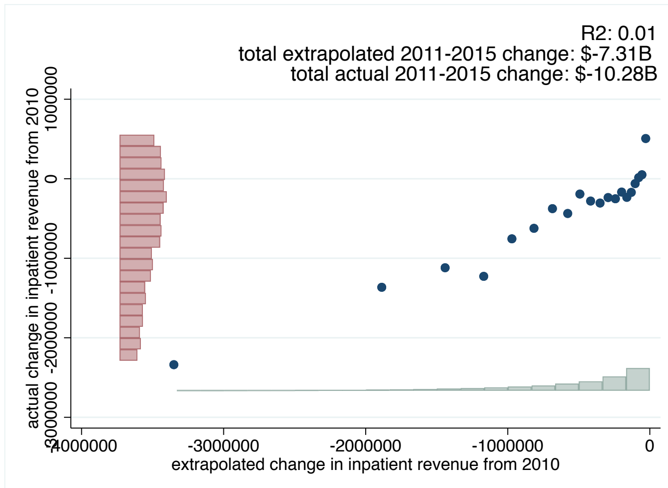
Figure E20. Marginal Value of Public Funds by Year, Additional Assumptions



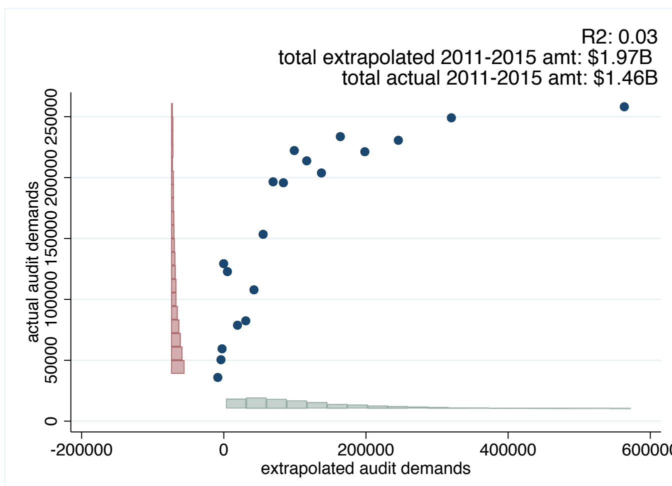
This figure plots the marginal value of public funds of a one percentage point increase in the RAC audit rate under different assumptions (a) on compliance costs and deterrence effect, (b) patient health effect, (c) treatment costs, and (d) Medicare revenue savings, as detailed in Section D.2 and summarized in Figure 8b.

Figure E21. Extrapolation Exercise: Actual vs. Extrapolated Savings

(a) Savings from changes in Medicare inpatient revenue



(b) Savings from audit demands



This figure plots binscatters of the actual versus extrapolated savings between 2011 and 2015 from (a) the reductions in Medicare inpatient revenue and (b) the payments demanded from audits. Actual changes in Medicare inpatient revenue are calculated by subtracting a hospital's revenue in a given year (between 2011 and 2015) from its 2010 revenue. Actual audit demands are calculated using the RAC audit data, and adjusted for refunds to hospitals due to the lawsuit over appeals described in Section A.2. Each observation is a hospital-year. Section E describes in further detail how the extrapolated changes in Medicare inpatient revenue and audit demands are calculated. The sample is winsorized at the 99th percentile of actual changes in Medicare inpatient revenue.

G Appendix Tables

Table F1. Summary statistics of 2010 hospital characteristics by sample

	(1) All Hospitals	(2) Border Hospitals
	A. Hospital Characteristics	
2011 audit rate	2.16	2.23
	(2.03)	(2.08)
Region A	0.17	0.08
Region B	0.19	0.36
Region C	0.42	0.37
Region D	0.22	0.18
Share urban	0.72	0.55
Share non-profit	0.63	0.70
Share for-profit	0.19	0.16
Share government	0.18	0.14
Beds	202.16	177.41
	(177.33)	(171.06)
Total cost (million $)	199.23	160.96
	(250.93)	(247.87)
Net admin costs (million $)	29.17	24.25
	(36.63)	(37.59)
	B. Medicare Inpatient Admission Characteristics	
Admissions	3465.75	3151.42
	(3205.86)	(3069.49)
Mean payment ($)	8617.36	7366.40
	(3179.31)	(2349.10)
Total payments (million $)	34.00	27.51
	(39.96)	(35.80)
Average short stay share	0.31	0.32
	(0.08)	(0.07)
Average circulatory diagnosis share	0.21	0.21
	(0.07)	(0.06)
Observations	2960	510

This table presents 2010 summary statistics of hospital characteristics and Medicare inpatient admissions in the overall and border samples. The border sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group. Standard deviation is in parentheses. Bed size, urban status, and profit type status come from the Medicare Provider of Services file. Total and administrative costs come from HCRIS. Medicare admissions and inpatient stay characteristics are from MEDPAR. Mean inpatient characteristics are defined as the average of each hospital's average (i.e., weighted by hospitals rather than claims). Short stay share is the share of Medicare admissions with length of stay ≤ 2. Circulatory diagnosis share is the share of Medicare admissions with a circulatory Major Diagnostic Category diagnosis.

Table F2. Correlation between 2010 hospital characteristics and 2011 audit rate

	(1) beds	(2) urban	(3) for profit	(4) non-chain	(5) total costs (millions)	(6) admin costs (millions)	(7) Medicare admissions	(8) inpatient revenue (millions)	(9) short stay share
					Panel A: Border Sample				
2011 audit rate	-3.82	-0.02**	-0.02	-0.00	1.53	-0.43	-120.16	-0.88	0.00*
	(4.33)	(0.01)	(0.01)	(0.01)	(5.66)	(0.70)	(71.29)	(0.70)	(0.00)
Nbr group FE	X	X	X	X	X	X	X	X	X
N Hosp	510	510	510	510	510	510	510	510	510
					Panel B: Overall Sample				
2011 audit rate	-12.82***	-0.02**	-0.03***	0.03***	-7.52*	-0.66	-241.67***	-2.38***	0.01***
	(2.93)	(0.01)	(0.01)	(0.01)	(3.79)	(0.62)	(50.90)	(0.52)	(0.00)
N Hosp	2960	2960	2960	2758	2960	2960	2960	2960	2960

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the state level. Panel A reports the coefficients from regressing the 2011 audit rate on an outcome variable in 2010 in the border sample, with neighbor comparison group fixed effects. The border sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group. Panel B reports the coefficients from regressing the 2011 audit rate on an outcome variable in 2010 in the overall sample. Bed size, urban status, and profit type status come from the Medicare Provider of Services file. Non-chain status comes from hospital merger data via Cooper et al. (2019). Total and administrative costs come from HCRIS. Medicare admissions and inpatient stay characteristics are from MEDPAR. Mean inpatient characteristics are defined as the average of each hospital's average (i.e., weighted by hospitals rather than claims). Short stay share is the share of Medicare admissions with length of stay ≤ 2.

Table F3. Summary Statistics of 2010 Inpatient Characteristics, by Sample

	(1)	(2)	(3)	(4)
	MEDPAR Sample		SID/SEDD Sample	
	All	*Border (100 mile)*	*FL ED*	*ED, 3 hr*
average age	73.04	73.35	74.10	72.59
	(14.03)	(13.66)	(14.19)	(15.12)
share female	0.56	0.56	0.55	0.54
	(0.50)	(0.50)	(0.50)	(0.50)
share white	0.82	0.87	0.83	0.81
	(0.39)	(0.33)	(0.38)	(0.39)
share inpatient last 30d	0.16	0.16	0.15	0.16
	(0.37)	(0.36)	(0.36)	(0.37)
Observations	11919671	2681021	602059	88027

This table presents 2010 summary statistics of traditional Medicare beneficiaries receiving inpatient stays in the following samples: all hospitals (column 1), hospitals within 100 miles of the border (column 2), patients admitted as inpatient from a Florida ED (column 3), and patients admitted as inpatient from a Florida ED who arrived at the ED within 3 hours of midnight (column 4). Data: MEDPAR and HCUP SID/SEDD.

Table F4. ED Arrival Hour Manipulation Tests

	(1)	(2)
	$[23{:}00 \leq T_v \leq 23{:}59]$	$\mathbb{1}[T_v \geq 00{:}00]$
$\mathbb{1}[q \geq 2013Q3]$	-0.001	-0.003
	(0.001)	(0.002)
Observations	1511606	1511606

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered by the ED arrival hour and quarter. This table reports estimates and standard errors of the coefficient on $\mathbb{1}[q \geq 2013Q3]$, an indicator for whether the ED visit occurred after the Two Midnights rule was implemented in 2013Q3. $[23{:}00 \leq T_v \leq 23{:}59]$ is an indicator equal to 1 if a patient's ED arrival hour is between 11:00PM and midnight, and 0 otherwise. $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether at patient's ED arrival hour was after midnight. Regression includes hospital fixed effects. Sample consists of traditional Medicare patients who arrived in the ED within 3 hours of midnight in a Florida hospital. Data: HCUP SID/SEDD.

Table F5. After-Midnight ED Arrival Coefficient on Stay Characteristics and Patient Outcomes

	(1) Total Charges ($)	(2) N Diagnoses	(3) N Procedures	(4) OR Procedure	(5) Revisit 60d	(6) Revisit 90d
β	42.707	-0.003	-0.005	-0.001	0.002	0.000
	(254.406)	(0.013)	(0.009)	(0.001)	(0.002)	(0.002)
Observations	1252735	1254857	1254857	1254857	1254857	1254857

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the ED arrival hour and quarter level. This table reports the β coefficient on $\mathbb{1}[q \geq 2013Q3] \times \mathbb{1}[T_v \geq 00{:}00]$ of the specification in Equation 7, where $\mathbb{1}[q \geq 2013Q3]$ is an indicator for whether the visit occurred after the Two Midnights rule was implemented in 2013Q3, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. "OR procedure" is an indicator for whether a patient received an OR procedure during their stay. "Revisit within 60/90 days" is an indicator for whether the patient had another ED visit or inpatient stay within 60/90 days of the ED visit. Sample comprises traditional Medicare patients who arrived in the ED within 3 hours of midnight in a Florida hospital. Regression includes hospital, hospital-quarter, hospital-hour fixed effects, and controls for age-sex bin, race, Hispanic indicator, point of origin indicator, last ED visit within 30 days indicator, number of chronic conditions, and zip code income. Data: HCUP SID/SEDD.

Table F6. Across-Hospital Post-2011 Coefficient

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall		LOS ≤ 2		Admin Costs	Software Installation
	Log Adm.	*Log Rev.*	*Log Adm.*	*Log Rev.*	*Log Costs*	*Medical Necc.*
2011 audit rate	-0.0154	-0.0166	-0.0227**	-0.0234***	0.0087	0.0153*
$\times$ post-2011	(0.0092)	(0.0136)	(0.0096)	(0.0056)	(0.0100)	(0.0081)
Hosp FE	X	X	X	X	X	X
Nbr group FE	X	X	X	X	X	X
Hosp	510	510	510	510	510	506
N	52139	52139	52139	46437	52107	36906
F	104.98	104.98	104.98	104.61	104.68	84.15

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the state and border segment level. This table reports the coefficients of the reduced form event study in Equation 5, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome after 2011. Columns 1-2 report two stage least squares outcomes for the number of and revenue from Medicare admissions overall, columns 3-4 report outcomes for the number of and revenue from Medicare admissions with length of stay ≤ 2, column 5 reports the outcomes for log net administration costs, and column 6 reports the outcomes for an indicator for installation of medical necessity software. Length of stay is counted as the difference in days between the admission and discharge date. Inpatient revenue is the sum of all Medicare inpatient payments. Net administration costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Indicator for installing software is equal to 1 if a hospital reports the status of a medical necessity software as "contracted/not yet installed," "installation in process," and "to be replaced" in HIMSS. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group.

Table F7. Heterogeneity of Across-Hospital Post-2011 Coefficient

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall		LOS $\leq$ 2		Admin Costs	Software Installation
	Log Adm.	*Log Rev.*	*Log Adm.*	*Log Rev.*	*Log Costs*	*Medical Necc.*
Panel A: Urban						
2011 audit rate × post-2011	-0.0410***	-0.0226	-0.0513***	-0.0215*	-0.0042	0.0130
	(0.0131)	(0.0145)	(0.0130)	(0.0113)	(0.0096)	(0.0082)
2011 audit rate × post × Urban	0.0367***	0.0086	0.0410***	-0.0017	0.0185**	0.0034
	(0.0090)	(0.0069)	(0.0109)	(0.0108)	(0.0083)	(0.0064)
Panel B: Teaching						
2011 audit rate × post-2011	-0.0195**	-0.0200	-0.0254**	-0.0235***	0.0042	0.0154
	(0.0082)	(0.0135)	(0.0105)	(0.0081)	(0.0104)	(0.0100)
2011 audit rate × post × Teaching	0.0195	0.0162	0.0131	0.0037	0.0217***	-0.0008
	(0.0131)	(0.0112)	(0.0177)	(0.0153)	(0.0069)	(0.0147)
Panel C: Hospital Profit Type						
2011 audit rate × post-2011	-0.0100	-0.0136	-0.0164*	-0.0199***	0.0116	0.0136*
	(0.0104)	(0.0143)	(0.0092)	(0.0069)	(0.0097)	(0.0073)
2011 audit rate × post × For-Profit	-0.0357*	-0.0386**	-0.0517**	-0.0539**	-0.0318	0.0169
	(0.0182)	(0.0162)	(0.0217)	(0.0256)	(0.0216)	(0.0114)
2011 audit rate × post × Gov't	-0.0258*	-0.0098	-0.0279	-0.0041	-0.0103	0.0030
	(0.0147)	(0.0130)	(0.0181)	(0.0178)	(0.0159)	(0.0075)
Panel D: Chain vs. non-chain						
2011 audit rate × post-2011	-0.0079	-0.0148	-0.0071	-0.0167*	0.0119	0.0193***
	(0.0140)	(0.0162)	(0.0110)	(0.0082)	(0.0094)	(0.0061)
2011 audit rate × post × Non-chain	-0.0150	-0.0037	-0.0312**	-0.0121	-0.0063	-0.0067
	(0.0122)	(0.0097)	(0.0143)	(0.0107)	(0.0044)	(0.0083)
Panel E: Bed Size						
2011 audit rate × post-2011	-0.0364***	-0.0260*	-0.0433***	-0.0231*	0.0015	0.0090
	(0.0104)	(0.0140)	(0.0126)	(0.0131)	(0.0110)	(0.0139)
2011 audit rate × post × Above Avg Beds	0.0419**	0.0187	0.0410**	0.0009	0.0144	0.0133
	(0.0165)	(0.0124)	(0.0173)	(0.0182)	(0.0090)	(0.0147)
Panel F: Medical Necessity Software Installed in 2010						
2011 audit rate × post-2011	-0.0172	-0.0210	-0.0188	-0.0204**	0.0187	0.0258***
	(0.0156)	(0.0177)	(0.0121)	(0.0093)	(0.0115)	(0.0051)
2011 audit rate × post × Med. Necc. App.	0.0035	0.0081	-0.0070	-0.0042	-0.0183	-0.0164***
	(0.0131)	(0.0103)	(0.0136)	(0.0099)	(0.0127)	(0.0051)
Hosp	510	510	510	510	510	506
N	52139	52139	52139	52118	52107	36906

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors are in parentheses and are clustered at the state and border segment level. This table reports the coefficients of the reduced form event study in Equation 5, scaled by the correlation between the leave-one-out 2011 audit rate and the actual 2011 audit rate in the weighted border hospital sample. The omitted year is 2010. Each coefficient represents the effect of a one percentage point increase in 2011 audit rate on a hospital-level outcome after 2011. Columns 1-2 report two stage least squares outcomes for the number of and revenue from Medicare admissions overall, columns 3-4 report outcomes for the number of and revenue from Medicare admissions with length of stay $\leq$ 2, column 5 reports the outcomes for log net administration costs, and column 6 reports the outcomes for an indicator for installation of medical necessity software. Length of stay is counted as the difference in days between the admission and discharge date. Inpatient revenue is the sum of all Medicare inpatient payments. Net administration costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Indicator for installing software is equal to 1 if a hospital reports the status of a medical necessity software as "contracted," "installation in process," and "to be replaced" in the HIMSS data in 2012. The sample comprises hospitals within a hundred miles of the RAC border with at least 1 hospital in their neighbor comparison group. Omitted year is 2010.

Table F8. After-Midnight ED Arrival Coefficient, Heterogeneity by Hospital Chars.

	(1)	(2)	(3)	(4)	(5)	(6)
			Inpatient			
β	0.011*	-0.005**	-0.004*	-0.008***	-0.007***	0.002
	(0.005)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)
× Urban	-0.019**					
	(0.005)					
× Teaching		-0.006*				
		(0.003)				
× For-profit			-0.007*			
			(0.003)			
× Gov't			-0.003			
			(0.006)			
× Non-chain				0.003		
				(0.006)		
× Above Avg. Beds					0.010**	
					(0.003)	
× Med. Necc. App						-0.013***
						(0.003)
Observations	1246862	1246856	1246862	1222485	1246862	1203528

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the ED arrival hour and quarter level. This table reports the β coefficient on $\mathbb{1}[q \geq 2013Q3] \times \mathbb{1}[T_v \geq 00{:}00]$ of the specification in Equation 7, interacted with hospital characteristics. $\mathbb{1}[q \geq 2013Q3]$ is an indicator for whether the visit occurred after the Two Midnights rule was implemented in 2013Q3, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. "Inpatient" is an indicator variable for whether the patient was eventually admitted as inpatient from the ED (HCUP SID/SEDD). The sample consists of traditional Medicare patients who arrived in the ED within 3 hours of midnight in a Florida hospital. Regression includes hospital, hospital-quarter, hospital-hour fixed effects, and controls for age-sex bin, race, Hispanic indicator, point of origin indicator, last ED visit within 30 days indicator, number of chronic conditions, and zip code income. Urban/rural, teaching/non-teaching, for-profit/government/non-profit, and bed size come form the Medicare Provider of Services file. Non-chain status come from Cooper et al. (2019). Medical necessity application is an indicator which is equal to one if medical necessity checking application is listed as "live and operational," "contracted" "installation in process," or "to be replaced" in the HIMSS data in 2012.

Table F9. Robustness Test: Sample of Patients by ED Arrival Relative to Midnight

	(1)	(2)	(3)	(4)	(5)
	Patient Sample				
	Within 1 Hour	Within 2 Hours	Within 3 Hours	Within 4 Hours	Within 5 Hours
	Panel A: Inpatient				
β	-0.007	-0.007**	-0.007***	-0.008***	-0.007***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
	Panel B: Revisit within 30 days				
β	-0.002	0.000	0.001	-0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Observations	394222	809058	1254857	1740915	2267496

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the ED arrival hour and quarter level. This table reports the β coefficient on $\mathbb{1}[q \geq 2013\text{Q3}] \times \mathbb{1}[T_v \geq 00{:}00]$ of the specification in Equation 7, where $\mathbb{1}[q \geq 2013\text{Q3}]$ is an indicator for whether the visit occurred after the Two Midnights rule was implemented in 2013Q3, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. Regression includes hospital, hospital-quarter, hospital-hour fixed effects, and controls for age-sex bin, race, Hispanic indicator, point of origin indicator, last ED visit within 30 days indicator, number of chronic conditions, and zip code income. The samples comprise of traditional Medicare patients who arrive at the ED in a Florida hospital within 1 hour of midnight (11PM-12:59AM; column 1), within 2 hours of midnight (10PM-1:59AM; column 2); within 3 hours of midnight (9PM-2:59AM; column 3); within 4 hours of midnight (8PM-3:59AM; column 4); and within 5 hours of midnight (7PM-4:59AM; column 5).

Table F10. After-Midnight ED Arrival Difference-in-Difference Coefficient, Heterogeneity by Patient Severity

	(1) Inpatient	(2) Revisit 30d
$\beta \times$ (Risk Decile 1)$_v$	0.015***	0.001
	(0.003)	(0.003)
$\beta \times$ (Risk Decile 2)$_v$	-0.006**	-0.002
	(0.002)	(0.005)
$\beta \times$ (Risk Decile 2)$_v$	-0.018***	0.001
	(0.004)	(0.005)
$\beta \times$ (Risk Decile 3)$_v$	-0.018***	0.009
	(0.007)	(0.006)
$\beta \times$ (Risk Decile 4)$_v$	-0.052***	0.004
	(0.008)	(0.006)
$\beta \times$ (Risk Decile 6)$_v$	-0.055***	-0.005
	(0.006)	(0.007)
$\beta \times$ (Risk Decile 7)$_v$	-0.036**	0.003
	(0.011)	(0.007
$\beta \times$ (Risk Decile 8)$_v$	-0.009	-0.008
	(0.014)	(0.005)
$\beta \times$ (Risk Decile 9)$_v$	-0.007	-0.000
	(0.010)	(0.004)
$\beta \times$ (Risk Decile 10)$_v$	-0.003	-0.002
	(0.004)	(0.005)
Observations	1236048	1236048

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors are in parentheses and are clustered by the ED arrival hour and quarter. This table reports the $\beta \times (\text{Risk Decile 1})_v$ coefficient on $\mathbb{1}[q \geq 2013Q3] \times \mathbb{1}[T_v \geq 00{:}00]$ of the specification in Equation 7, interacted with an indicator for the predicted risk decile of visit v. $\mathbb{1}[q \geq 2013Q3]$ is an indicator for whether the visit occurred after the Two Midnights rule was implemented in 2013Q3, and $\mathbb{1}[T_v \geq 00{:}00]$ is an indicator for whether the ED arrival hour for the visit was after midnight. Patient risk is predicted by estimating a logit using ED visits between 9AM and 3PM of an indicator for being admitted within 30 days of an ED visit on patient demographics, current ED visit information, and information on any prior visits in the last 365 days. Demographics include age-bin, sex, race, Hispanic indicator, point of origin indicator, and mean zip code income. Information on current visit includes hospital and quarter. Information on previous visits includes the number of visits/inpatient stays/length of stay in the last month or last year, as well as any diagnoses and procedures recorded in stays within the last month or last year.