The Costs and Benefits of Monitoring: Evidence from Medicare Audits

Maggie Shi *

Preliminary work in progress; do not circulate or cite.

(Click for latest version)

July 9, 2021

Abstract

Governments often outsource service delivery to third parties and subsequently monitor the third parties to deter unwanted behavior or spending. This paper studies monitoring in one of the largest federal programs, Medicare, where the third party – a healthcare provider – has considerable autonomy and makes decisions that can affect patient health. I find that monitoring providers decreases unnecessary government expenditure without affecting patient health, but that these savings also come at the expense of substantial provider compliance costs. I use novel data on a Medicare audit program which monitored and then reclaimed $8.4 billion in hospital inpatient revenue in its first five years (one percent of revenue at risk of audit). Hospitals respond to monitoring by reducing admissions, and fine-tune the response to not harm patients by targeting admissions that are more likely to be unnecessary. In order to comply with monitoring, hospitals incur compliance costs and adopt technology to identify unnecessary admissions. For every $1000 saved through monitoring, hospitals spent $173 in compliance costs. Despite the compliance costs, I find that increased monitoring improves welfare because of the substantial government savings.

*Department of Economics, Columbia University. I thank Wojciech Kopczuk, Adam Sacarny, Pietro Tebaldi, and Michael Best for their input and support with this project. I also thank Jetson Leder-Luis, Ajin Lee, Jon Skinner, Daniel Rees, and participants at EEA, APPAM Student Research Series, ASHEcon, WEAI Graduate Student Workshop, Columbia Health Policy, and Columbia Applied Micro 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). All errors are my own.

1 Introduction

Governments often contract with third parties to deliver services. As these third parties often have more information than the government as well as private incentives to maximize compensation, a key challenge for policymakers is keep spending down, conditional on the quantity or quality of services. In the context of public healthcare programs, this issue can arise when the third parties – healthcare providers – provide unnecessary services that increase their own compensation but have little to no effect (or even a negative effect) on patient health.[1] In the U.S., the Medicare program spends an estimated $13 billion on unnecessary healthcare each year (U.S. Department of Health and Human Services, 2019; Centers for Medicare and Medicaid Services, 2019). More broadly, it is estimated that 20 percent of healthcare expenditure goes towards wasteful spending (Berwick and Hackbarth, 2012). Given the efficiency cost of raising revenue and the opportunity cost of public funds that could go toward other programs, reducing wasteful expenditure is a longstanding goal for policymakers.

One potential solution for the government is to monitor the third party it contracts with to ensure compliance and cost-effectiveness (Holmström, 1979). For example, Medicare could monitor unnecessary service provision and either refuse to pay providers for it or even penalize providers. However, because healthcare delivery is complex and providers make care decisions *ex ante* (Arrow, 1963), it is unclear how providers would respond to monitoring – depending on how providers respond, monitoring may not improve net *welfare*, even if it saves Medicare money. In practice, the savings derived from additional monitoring may not be welfare-improving from a societal point of view if monitoring imposes high costs, or if it changes provider behavior in a way that harms patients. Thus to understand whether increasing monitoring improves welfare, we need to know the magnitudes of the savings and costs associated with monitoring healthcare providers.

This paper studies how healthcare providers respond to monitoring, and in turn how these responses factor into government savings, provider compliance costs, and patient health outcomes. I focus on a large monitoring program within Medicare: the Recovery Audit Contractor (RAC) program. The Centers for Medicare and Medicaid Services (CMS) pay for services for Medicare beneficiaries by contracting with healthcare providers, who deliver care and then submit claims to CMS for reimbursement. RAC auditors then select individual

[1] I use a narrow notion of the term "unnecessary," referring to CMS's definition of medically necessary care rather circularly as "services or supplies that: are proper and needed for the diagnosis or treatment of [the] medical condition, are provided for the diagnosis, direct care, and treatment of [the] medical condition, meet the standards of good medical practice in the local area, and aren't mainly for the convenience of [the patient] or [his] doctor" (Centers for Medicare and Medicaid Services, 2006).

claims to examine more closely for payment errors such as overpayments for unnecessary services. Since their inception in 2010, RACs have conducted over 18.7 million audits and reclaimed $18 billion in erroneous payments across all programs within Medicare that it monitors.

I focus on RAC audits of inpatient hospital admissions to understand the savings, costs, and patient health effects of monitoring healthcare providers. In the first five years of the program, RACs audited four percent of Medicare inpatient admissions; on net, RACs reclaimed $8.4 billion of inpatient revenue (about one percent of the $910 billion of inpatient revenue at risk of audit), making it the largest source of reclaimed payments in the RAC program.[2] To study the RAC program, I use novel, audit-level data acquired through a Freedom of Information Act request on 4.5 million inpatient audits. I then link this to data on Medicare claims, hospital costs, hospital technology adoption, emergency department and inpatient discharges, and patient outcomes.

I first identify the effect of RAC audits on provider-level outcomes such as hospital revenue, compliance costs, and technology adoption by exploiting plausibly exogenous variation in audit rates across hospitals. This variation is driven by the audit intensity of the particular RAC assigned to a hospital's region. Comparing neighboring hospitals on opposite sides of the border between different firms' regions, I find that increased auditing substantially reduces Medicare expenditure, and these savings are driven by reductions in unnecessary admissions. RAC audits mostly deter admissions that CMS is more likely to consider unnecessary – namely, short hospital stays and stays for diagnoses more prone to payment errors for lack of medical necessity. Overall, increasing a hospital's audit rate by one percentage point (46 percent) reduces payments for future admissions by 2.4 percent, *on top* of reclaimed payments from audit findings. The reductions in admissions persists even in later years, when the RAC program is significantly scaled back.

But in order to achieve these savings, hospitals incur significant compliance costs. A one percentage point increase in audit rate leads to an immediate 1.5 percent increase in hospital administration costs which subsides in subsequent years. For every $1000 in Medicare savings, hospitals spent $173 in compliance costs. One source of these compliance costs is the technology that hospitals adopt to identify unnecessary admissions and prevent future audits – hospitals subject to higher audit rates install software specifically aimed at identifying unnecessary care. The immediate increase in costs, persistent decrease in admissions, and uptake of technology to detect unnecessary admissions suggest that the higher audit rates induced hospitals to make investments in order to identify unnecessary expenditure

[2] For comparison, the Hospital Readmission Reduction Program levied $1.9 billion in penalties to hospitals in its first five years (Boccuti and Casillas, 2017).

and mitigate future audits. In short, in response to the threat of *ex post* audit, hospitals spent money to invest in technology to avoid the audits *ex ante*.

We may also worry that hospitals that reduce admissions will unintentionally deny care to a patient who actually needs to be admitted; this could in turn harm patient health outcomes. But since we find that audits reduce hospital admissions, this means that a change in a hospital's average patient health could be driven by a combination of the causal effect of monitoring *and* of changing patient composition. Motivated by this, I turn to a separate, patient-level empirical strategy to study health outcomes. This within-hospital strategy holds patient composition fixed by comparing patients in the same hospital who are subject to different audit likelihoods. I leverage a policy which varied a patient's audit likelihood depending on a factor outside of a hospital's control: the hour at which a patient first arrives at the emergency department (ED). In particular, the policy increased audit likelihoods for patients who arrived at the ED *after* midnight. Accordingly, hospitals cut back on admissions for after-midnight ED patients. But I do not find evidence that patients denied admission due to the midnight policy were more likely to revisit the hospital in the next 30 days. These findings suggest that hospitals responded to monitoring by reducing unnecessary admissions but not necessary ones, perhaps as a result of adopting technology to identify unnecessary admissions.

Overall, the empirical results demonstrate that a marginal increase in monitoring reduces unnecessary Medicare expenditure; however, these savings come at a substantial private compliance cost to providers. The tradeoff between government savings and private compliance costs motivates a comparison of the relative magnitudes of each effect on total welfare. I incorporate the empirical estimates into a sufficient statistics framework that accounts for government savings, provider compliance costs, government administrative costs, and patient health effects. Despite substantial compliance costs, increasing the audit rate is welfare-improving over a 5-year horizon. But the size of this welfare effect depends crucially on provider compliance costs – the welfare improvement would be *7.1 times* larger if hospitals did not incur any compliance costs.

This paper contributes to the health policy, health economics, and public economics literatures. I contribute to the health policy literature with, to my knowledge, the first quasi-experimental evidence on the impact of monitoring in Medicare. Despite the large impact of the RAC program and its counterparts, there is little academic work studying it, aside from qualitative studies of select hospitals (Sheehy et al., 2015, 2017). I contribute to our understanding of these monitoring programs by highlighting costs and savings that go beyond what policymakers consider in their cost-benefit analyses (Centers for Medicare and Medicaid Services, 2011). While policymakers only considered the total payments corrected by RACs, I find that this calculation is missing two important factors: the savings from reductions in unnecessary care and the costs for providers to comply.

This paper also contributes to the health economics literature by shedding light on how providers respond to insurers' non-financial, administrative actions. Administrative mechanisms like monitoring reduce the *realized* price for care, but only for relatively low-severity cases. It is well-established that healthcare providers respond to *contracted* prices, by either changing the quantity and type of care provided (Cutler, 1995; Ellis and McGuire, 1996; Clemens and Gottlieb, 2014; Alexander and Schnell, 2019; Gross et al., 2021; Gupta, 2021), or by changing how they document the care they provide (Silverman and Skinner, 2004; Dafny, 2005; Sacarny, 2018). In contrast, we know much less about how providers respond to administrative mechanisms that change the realized price, despite the fact that such mechanisms are used widely by almost all insurers, both public and private. By studying provider responses to systematic monitoring, I contribute to a more-recent literature on administrative mechanisms like billing complexity (Dunn et al., 2021), fraud detection (Leder-Luis, 2020; Howard and McCarthy, 2021), and prior authorization (Burn et al., 2021; Roberts et al., 2021). I find that monitoring providers encourages them to reduce expenditure on low-value care, but at the cost of adding to providers' already-large administrative burden (Cutler and Ly, 2011; Himmelstein et al., 2014; Papanicolas et al., 2018; Dunn et al., 2020).

More generally, I contribute to the public economics literature by highlighting a potential downside to well-intentioned public policy: high compliance costs for the third parties that interact with the government. While monitoring deters wasteful public expenditure without affecting quality of services provided, I also find evidence that it leads to considerable compliance costs for the providers subject to monitoring. This empirical finding underlines the theoretical point made by Keen and Slemrod (2017), that simply reducing government expenditure (or recovering government revenue) is not sufficient for a policy to be welfare-improving; it must also not be associated with high compliance or administrative costs. Previous work on other public programs has found that third parties often face private costs when they interact with the government for their own gain – for example, when applying for benefits (Besley and Coate, 1992; Currie, 2006; Deshpande and Li, 2019) or requesting tax refunds and credits (Kopczuk and Pop-Eleches, 2007; Zwick, 2021). I document an instance where third parties incurred substantial private costs to *save* money on behalf of the government, going so far as to install technology to identify wasteful expenditure. The marginal welfare improvement from increasing monitoring would be 7.1 times greater absent these costs. This highlights the importance of taking compliance costs into account when evaluating other contexts in which policymakers aim to recover tax revenue or reduce government expenditure, such as tax audits or Social Security disability examinations.

2 Policy Context

This paper focuses on the component of Medicare that pays for hospital admissions, which comprises the largest service expenditure category for the Medicare program – CMS spent $147 billion (19 percent of its total expenditure) on inpatient hospital admissions in 2019 (Medicare Payment Advisory Commission, 2020). CMS pays for inpatient hospital admissions through a prospective payment system in which CMS reimburses hospitals a fixed payment per inpatient stay, within broad categories of diagnoses. Hospitals keep the difference between the fixed payment and the costs to treat the patient, so they have incentive to keep costs low within an admission. The payment rate depends mostly on the patient's diagnosis and pre-existing health conditions – importantly, it typically does *not* depend on characteristics of the inpatient stay itself, like length of stay (LOS).

In the early 2000s, policymakers became increasingly concerned with one area of vulnerability: unnecessary inpatient stays, which were particularly common among short 0-2 day stays (Centers for Medicare and Medicaid Services, 2011). The Medicare Payment Advisory Commission contended that hospitals were admitting patients for short stays, even when they were unnecessary, because short stays are very profitable. As the payment rates for hospital stays are set based on averages across *all* stays within a diagnosis category, stays with below-average costs are more profitable. In 2012, the average payment-to-cost ratio was 1.55 for one-day inpatient stays and 1.3 for two-day stays, whereas it was 0.72 for stays eight days or longer (Medicare Payment Advisory Commission, 2015).

To address these unnecessary inpatient stays, in 2011 CMS directed Recovery Audit Contractors (RAC) firms to begin monitoring and reclaiming payments for them. The program was conducted through contractor firms, each of which was responsible for auditing claims in the last three years in a geographic region. The regions are demarcated by the red line in Figure 2a. For most inpatient claims, a RAC audit consists of the firm running a proprietary algorithm on claims data to flag individual claims for potential issues like missing documentation, incorrect coding, or unnecessary services. A medical professional (e.g., a nurse or a coder) then requests and reviews all documentation associated with the flagged claim. The medical professional then makes a determination about whether CMS made an overpayment (or, for a small percentage of claims, an underpayment). If the RAC determines that there was an overpayment, then CMS corrects it by demanding a payment back from the provider. There is no additional penalty associated with a corrected payment. The firms are paid a negotiated percentage of the payments they correct (9–12.5 percent), after appeals.

Figure 1 illustrates the value of audited hospital stays and reclaimed payments from these audits for the average hospital, by year of audit. RACs earned the majority of their revenue from auditing inpatient admissions, as inpatient admissions make up the largest share of erroneous payments within Medicare (Centers for Medicare and Medicaid Services, 2011). The program was large and affected almost every hospital – 96 percent of short-term acute hospitals were audited by RACs. There is a substantial increase in auditing activity in 2011, corresponding to the expansion of RAC's auditing scope to include unnecessary admissions. At the program's peak in 2012, RACs reclaimed an average of $1 million per hospital. Given that the average hospital earned $32 million in Medicare inpatient revenue in 2012, RAC auditing had an economically meaningful impact on hospitals' bottom line.

3 Empirical Strategies and Data

Given the size and scope of the RAC program, one natural question is how hospitals responded to it, and consequently how these responses translated into Medicare savings, hospital compliance costs, and patient outcomes. Medicare savings could come from the payments demanded back from prior admissions as well as money not spent on future admissions deterred by audits. Hospitals could incur compliance costs from both the day-to-day cost of managing audit requests and from one-time investments made to mitigate future audits. Additionally, if hospitals changed the amount of care provided in response to audits, then this could also affect patient health outcomes.

To investigate these potential effects, I deploy two separate empirical strategies: one across hospitals and one within hospitals. The across-hospital strategy allows me to study hospital-level outcomes like admissions, revenues, and costs across hospitals with plausibly exogenously different audit rates. I then switch to a within-hospital strategy to study patient-level health outcomes. I make this switch because if hospitals change the number of inpatient admissions in response to RAC audits, then each hospital's patient composition will mechanically change as well. With an across-hospital comparison, it is difficult to separate these patient composition effects from the causal effect of monitoring. Motivated by this concern, I use a within-hospital strategy that compares patients within the same hospital who have plausibly exogenously different audit likelihoods.

The across-hospital empirical strategy leverages the sharp change in audit intensity across the borders that separate different RAC firms' regions. I focus on hospitals close to the border and instrument for a hospital's 2011 audit rate with a leave-one-out jackknife audit rate of other hospitals subject to the same RAC firm. I incorporate the instrumented audit rate into a dynamic difference-in-difference framework comparing neighboring hospitals along the border, before and after the expansion of auditing scope in 2011.

The within-hospital empirical strategy compares patients with different audit likelihoods in a setting with fewer concerns about patient composition effects – patients who arrive at the emergency department (ED). This strategy is based on the Two Midnights rule, which stated that if a patient spent two or more *midnights* in the hospital (including time in the ED), they could not be audited for medical necessity. Because of this rule, patients who arrived in the ED just before midnight had a mechanically lower audit likelihood than patients who arrived just after midnight. I exploit this change in audit likelihood at midnight with a difference-in-difference strategy that compares Medicare patients who arrive at the ED before vs. after midnight, pre- and post-Two Midnights rule. I next describe the empirical strategies and relevant data in further detail.

3.1 Across-Hospital Empirical Strategy and Data

Empirical Strategy In the across-hospital empirical strategy, I use a border difference-in-difference strategy to exploit the variation in the identity, and thus audit intensity, of a hospital's RAC for hospitals along the border between RAC regions. I compare hospitals to their neighbors on the other side of the border, who are subject to either a more or less aggressive RAC. Figure 2a illustrates the variation in audit rate at the border between RAC regions. It plots the 2011 average audit rate for each state, with a darker shade corresponding to higher audit rate. There are sharp changes in audit rate when going from one side of the RAC region border to the other – for example, the audit rate is much higher in Region D compared to Region C.

In order to proxy for the audit intensity of a hospital's RAC without capturing variation driven by unobserved hospital-specific factors, I instrument for a hospital's audit rate with its jackknife state audit rate, defined as the leave-one-out average audit rate of other hospitals in the same state. The jackknife state audit rate captures spillovers from hospitals that share a RAC with an individual hospital via the shared RAC algorithm used for flagging claims. Even if a hospital itself has very few erroneous payments, a RAC may be less willing to give it the "benefit of the doubt" if many of its neighbors have high levels of errors, as the algorithm is tailored to the average hospital in a given region. This instrument also captures any state policies that may influence how aggressive a RAC is in a particular state (e.g., state Medicaid policies, tort laws).

Using the instrumented audit rate, I compare a border hospital to its neighboring hospitals on the other side of the border – each hospital is compared to a unique set of geographically-close neighbors. The main results are based on defining border hospitals as all hospitals within 100 miles of the RAC border, and neighboring comparison hospitals as all hospitals within 100 miles of a given hospital, on the other side of the border. The results are robust to specifications using varying distances.

Identifying the causal effect of a higher 2011 audit rate requires the parallel trends assumption, that absent the introduction of RAC audits, hospitals on the high side of the border have trended similarly to their neighbors on the low side. We can check this assumption by plotting the coefficients from a dynamic specification and checking for pre-trends before 2011. I use a local comparison among border hospitals that are geographically close to each other to address the concern that, if we compared all hospitals in high-audit regions to all hospitals in low-audit regions, hospitals in different regions may be subject to geographically-correlated trends over time. For example, we may be worried that different regions experience differential healthcare expenditure trends (Finkelstein et al., 2016; Skinner, 2011), which could violate the parallel trends assumption. Comparing hospitals with above- and below-median audit rates within the overall sample and the border hospital sample (Table 1), we see that the border hospital sample is relatively more balanced in pre-reform observables – for example, the difference in Medicare admissions volume between above- and below-median border hospitals is smaller than the difference in the overall sample. This suggests that hospitals in different regions along the RAC border are more comparable to each other than hospitals in the overall sample.

Equations 1 and 2 define the two stages of the two stage least squares regression specification for the across-hospital strategy.

First Stage:

$$\begin{aligned} Audit\ Rate_{h,2011} \times Year_y = &\gamma_1 Jackknife\ Audit\ Rate_{hs(h),2011} \times Year_y \\ &+ \alpha_h + \lambda_{g(h)} \times Year_y + \nu_{ht} \end{aligned} \tag{1}$$

Second Stage:

$$\begin{aligned} Y_{ht} = &\beta_0 + \sum_{y=2007}^{2015} \beta_{1y} \widehat{Audit\ Rate}_{h,2011} \times Year_y \\ &+ \alpha_h + \lambda_{g(h)} \times Year_y + \varepsilon_{ht} \end{aligned} \tag{2}$$

The first stage of the two stage least squares specification (Equation 1) predicts hospital h's 2011 audit rate using its jackknife 2011 state audit rate (interacted with a year dummy), where hospital h is in state $s(h)$. The second stage incorporates the hospital's predicted 2011 audit rate into a dynamic difference-in-difference specification. The specification includes hospital fixed effects α_h and neighboring hospital comparison group-year fixed effects $\lambda_{g(h)} \times Year_y$. Y_{ht} is the outcome of interest for hospital h in year t, such as the number of inpatient admissions or amount of administrative spending. 2010 is the omitted year. The coefficient

β_{1y} can be interpreted as the effect of a one percentage point increase in 2011 audit rate on Y_{ht}, relative to its 2010 level. The estimates are clustered at the state and 100-mile border segment level. Following other papers using border difference-in-difference strategies, I cluster at the border segment level, which accounts for repeated hospital observations (Dube et al., 2010).

Data The across-hospital empirical strategy uses four main datasets. First, I use novel administrative data from the Recovery Audit Contractor program acquired through a FOIA request. This data spans 2010 to 2020, and includes claim-specific information on every single audit RACs conducted. This data covers 4.5 million audits of inpatient stays. Second, I merge this with Medicare inpatient claims data (Medicare Provider Analysis and Review; MEDPAR) and outpatient claims data between 2007 and 2015. Third, I measure hospital costs data from the Healthcare Cost Report Information System (HCRIS), which collects cost reports that hospitals submit to CMS. 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. I also use the Medicare Provider of Services file and hospital merger data from Cooper et al. (2019) to study hospital characteristics.

3.2 Within-Hospital Empirical Strategy and Data

Empirical Strategy I next turn to the within-hospital, patient-level empirical strategy which leverages the Two Midnights rule. I first explore whether having a higher audit likelihood affects how likely a patient is to be admitted from the ED. In 2013, 73 percent of Medicare inpatient admissions originated in the ED and 42 percent of Medicare ED visits result in an admission. If patients are denied admission because of a higher audit likelihood, I then trace out whether it affected patient health outcomes. The logic is that if providers deny admission to avoid auditing, and this reduces quality of care, then we should see that outcomes for patients with higher audit likelihoods should worsen.

This strategy is based on the "Two Midnights rule," a CMS policy introduced in August 2013 which clarified when an inpatient admission was medically necessary, and thus which inpatient claims RACs were allowed to audit for medical necessity. Under this rule, CMS counted the number of *midnights* a patient's stay in the hospital crossed – including the time spent in the ED, in observation, and in inpatient. If the patient's stay crossed two midnights, then RACs were barred from auditing the stay for medical necessity. If the patient's stay did not cross two midnights, then RACs could audit it. CMS's arbitrary choice to count midnights created a discontinuity in audit likelihood depending on whether a patient arrived at the hospital before or after midnight.

I study whether the Two Midnights rule affected patient admissions and outcomes with a difference-in-difference strategy, dividing patients by whether they arrived before or after midnight and comparing them pre- and post- policy implementation in August 2013. Table 3 reports summary statistics for before- and after-midnight patients; after-midnight patients tend to be sicker, as they are more likely to be admitted, have more chronic conditions, and are more likely to have a recent hospital visit or inpatient stay. I focus on patients arriving in the ED within three hours of midnight. "Treated" patients are ED patients whose ED arrival hour (*not* their inpatient admission hour) is between midnight and 3AM, while "control" patients are ED patients whose ED arrival hour is between 9PM and midnight. I then compare the difference between these two sets of patients pre- and post-implementation in August 2013.

Equation 3 defines the main difference-in-difference specification for the within-hospital strategy.

$$
\begin{aligned}
Y_i = \beta_0 &+ \beta_1 (After\ MN)_{a(i)} (Post\ Aug\ 2013)_{t(i)} \\
&+ Hosp_{h(i)} \times Qtr_{t(i)} + Hosp_{h(i)} \times (ED\ Arrival\ Hr_{a(i)}) \\
&+ \boldsymbol{\gamma' X_i} + \varepsilon_i
\end{aligned} \tag{3}
$$

Y_i is the outcome of interest describing patient i's visit, such as a dummy for whether the patient was admitted from the ED or whether the patient revisited a hospital within a month of this visit. $(After\ MN)_{a(i)}$ is a dummy for whether patient i's ED arrival hour $a(i)$ was after midnight. $(Post\ Aug\ 2013)_{t(i)}$ is a dummy for whether the visit was pre- or post-policy implementation in August 2013. I also include hospital fixed effects, quarter fixed effects, and ED arrival hour fixed effects and their interactions, as well as controls for the following patient characteristics: patient age, race, Hispanic, point of origin, dummy for whether last ED visit was within three days, number of chronic conditions, and average income in patient's zip code. β_1 is the coefficient of interest, and can be interpreted as the effect of increased audit likelihood (from the Two Midnights rule) on after-midnight, post-2013 ED visits.

Interpreting β_1 as the causal effect of increased audit likelihood 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 throughout this time period. We can plot quarter-by-quarter β_{1t}'s in the dynamic difference-in-difference specification (Equation 4) to see if the coefficients before August 2013 are insignificant.

$$
\begin{aligned}
Y_i = \beta_0 + \beta_{1t} \sum_{t=2010Q1}^{2016Q4} (After\ MN)_{a(i)} Qtr_{t(i)} \\
+ Hosp_{h(i)} + Qtr_{t(i)} + (ED\ Arrival\ Hr_{a(i)}) \\
+ Hosp_{h(i)} \times Qtr_{t(i)} + Hosp_{h(i)} \times (ED\ Arrival\ Hr_{a(i)}) \\
+ \boldsymbol{\gamma}' \boldsymbol{X_i} + \varepsilon_i
\end{aligned} \tag{4}
$$

The second assumption required for identification is one which is often used in discontinuity empirical strategies – that there is no manipulation of the ED arrival hour. We require that hospitals are not mis-reporting ED arrival hour as a result of the Two Midnights rule, otherwise the comparison of before- and after-midnight patients after policy implementation is not equivalent to the same comparison before implementation. I check this empirically by looking at how the share of after-midnight patients has changed post-policy, among cohorts where ED arrival hour is more or less easy to manipulate. I also argue that, practically speaking, it is not easy for hospitals to manipulate ED arrival hour in response to this policy. The ED arrival hour is recorded as soon as the patient walks in the hospital, which makes it more difficult to manipulate. Additionally, in order to game the Two Midnights rule, providers would have to make after-midnight arrivals look like before-midnight arrivals. This kind of manipulation requires actively misreporting a patient's ED arrival time to an earlier time, rather than a more passive policy of "dragging their feet" in order to record a later arrival time.

Data To implement the empirical strategy, I use the Florida State Emergency Department Database (SEDD) and State Inpatient Database (SID), available through the Healthcare Cost and Utilization Project (HCUP) between 2010 and 2015. I focus on Florida because it is the only state that reports ED arrival hour – even Medicare's Inpatient and Outpatient files do not report this variable.

4 Results

4.1 Across-Hospital Results

In order to implement the across-hospital instrumented difference-in-difference strategy in Equations 1 and 2, I first construct the sample of border hospitals. For each hospital within 100 miles of the RAC border, I identify all hospitals within a 100 mile radius on the other side of the border, and call this the "neighboring hospital comparison group." Figure A3 illustrates an example of how a neighboring hospital comparison group is constructed. I then restrict the sample to hospitals with at least one hospital in its comparison group. Table 1 reports pre-reform summary statistics for hospitals with above- and below- median 2011 audit rates among the overall hospital sample and the border hospital sample. Figure A6 plots admission trends over time among border hospitals, split by RAC region. Overall, Medicare admissions are decreasing between 2007 and 2015.

First Stage To implement the across-hospital strategy, I next investigate the relevance of the instrument in Equation 1 in this sample. For each hospital h, I calculate its 2011 jackknife state audit rate $Jackknife\ Audit\ Rate_{hs(h),2011}$, where hospital h is state $s(h)$. Figure 3a plots a binscatter of the cross-sectional relationship between 2011 jackknife audit rate and 2011 hospital audit rate. In the border hospital sample, $Jackknife\ Audit\ Rate_{hs(h),2011}$ explains 34 percent of the variation in $Audit\ Rate_{h,2011}$. The binscatter shows that there is a linear relationship between the two, and the relationship is not driven by outliers. Figure 3b plots the coefficients from the dynamic instrumented difference-in-difference (Equation 2) on the amount of demanded payments from audited claims. A one percentage point increase in audit rate in 2011 is associated with \$314,115 in demands in 2011 per hospital. A higher audit rate in 2011 also results in additional reclaimed payments in subsequent years, although the magnitude diminishes over time. We see this pattern because audit rates are correlated across years (Figure A5) – one possibility could be that RACs face a fixed cost of starting audits at a particular hospital, for example the fixed cost of opening up channels of communication between the RAC and hospital administrators. Figure A2 shows that there is bunching at the audit rate of zero, consistent with there being a fixed cost of beginning audits at a hospital. The relationship between 2011 audits and audits in subsequent years means that we should interpret the results not just as a hospital's response to an increase in audits in 2011, but rather as the response to an increase in 2011 *and* anticipated increases in subsequent years. This also means that in evaluating the welfare effect of increasing audits in 2011, we need to take into consideration costs and savings not just in 2011, but over a longer multi-year horizon.

Admissions and Revenue Figure 4 presents the first set of main results – the Equation 2 coefficients on Medicare admissions and revenue outcomes. Table 2 column 1 reports the coefficients of the first stage and the yearly coefficients from 2011 to 2015. The coefficients represent the effect of a one percentage point increase in audit rate on the outcome of interest. 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. There is a lack of pre-trends prior to 2011 – hospitals with higher audit rates do not seem to be on differential trends than their neighboring hospitals with lower audit rates. This supports the parallel trends assumption that absent RAC audits, hospitals on different sides of the border would have trended similarly in the post-reform period.

Starting in 2011, there is a decline and then a plateau in admissions among hospitals with higher audit rates. A one percentage point increase in audit rate results in an average 1.72 percent decrease in admissions in 2012–2015. The steady decline and then plateau, rather than a sharp drop, likely reflects two factors: first, some of the 2011 admissions occurred before hospitals knew how aggressively they would be audited by RACs that year; and second, it may have taken time to implement practices to reduce unnecessary admissions in response to RAC audits. Inpatient revenue follows a similar trend, with an average 2.4 percent decrease in 2012–2015 relative to 2010. The coefficients for revenue are noisier because the actual revenue per admission is adjusted by patient- and hospital-specific factors. Despite being noisier, the coefficient on inpatient revenue is either a similar or greater magnitude than the coefficient on admissions. This suggests that the average hospital did not change its mean payment per stay, either by changing patient composition (i.e., denying admission mostly for stays with low DRG weights) or by upcoding (i.e., changing documentation practices to maximize payment).

I next focus on two types of admissions which are more likely to be unnecessary – short stay admissions and circulatory diagnosis admissions. CMS and other policymakers noted that unnecessary stays were more common among short stays (US Department of Health and Human Services Office of Inspector General, 2013; Miller, 2015). RACs' auditing behavior reflects this – 64 percent of RAC audits of inpatient stays in 2011 were for short stays, while only 29 percent of inpatient stays in 2011 were short stays. Figures A7a and A7b present coefficients of the log number of admissions with length of stay 0–2 and log revenue from these admissions. A one percentage point increase in audit rate results in an average 2.97 percent decrease in short stay admissions and a 3.85 percent decrease in revenue from short stays in 2012 and after. The year-by-year effects on short stays show a decline in 2011 that reaches its lowest point in 2012 before bouncing back slightly, indicating that hospitals may have "overcorrected" at first in cutting back on short stays. The large effects on short stays suggest that part of the overall decreases in short stays in this period can be attributed to RAC audits (Medicare Payment Advisory Commission, 2019). I then turn to admissions with a circulatory admissions, which are over-represented among unnecessary stays – three out of the five top service types with the highest payment error rates are associated with circulatory diagnoses (Centers for Medicare and Medicaid Services, 2011). A one percentage point increase in audit rate results in an average of 4.14 percent decrease in admissions with circulatory diagnoses and a 5.09 percent decrease in revenue from them in 2012 and after. Table B3 reports coefficients for admissions with length of stay greater than 2 and non-circulatory diagnoses; the effect of auditing on each is smaller but still negative. The larger responses for short stay admissions and circulatory admissions demonstrates that hospitals targeted their reductions among the types of claims that were more likely to be considered unnecessary.

The results on inpatient admissions and revenue demonstrate that hospitals responded to monitoring by reducing inpatient admissions, which led to substantial savings for Medicare. But these savings from fewer inpatient admissions could be mitigated if hospitals substituted away from inpatient admissions toward more outpatient care – for example to observation stays. Observation stays consist of short-term (often diagnostic) services provider at the hospital while a physician decides whether to admit a patient or send them home. Observation stays typically last less than 48 hours and are billed as an outpatient service, and are often cited as a more cost-effective alternative to a short inpatient stay (Medicare Payment Advisory Commission, 2015).[3] Figure A8 explores this by looking at the effects on log observation stays and log outpatient revenue. The results suggesting that at the hospital-wide level, there is little evidence that hospitals increased outpatient spending and observation stays in response to auditing. Thus, the savings from fewer inpatient admissions were not mitigated by substitution from inpatient care to outpatient care.

Compliance Costs and IT Adoption I next turn to measures of the administrative burden of RAC audits and compliance costs for providers. 75 percent of hospitals surveyed by the AHA reported increased administrative burden because of RAC auditing in 2012; the most commonly cited reasons were increased administration costs, training and education, and tracking software (Figure A9. Figure 4 and Table 2 columns 5-6 present results on two dimensions of administrative burden: administration costs and IT adoption. Figure 4c plots estimates of the effect of auditing on log administration costs as reported in annual Hospital Cost Reports that hospitals submit to CMS. A one percentage point increase in RAC auditing in 2011 results in a 1.5 percent uptick in administration costs immediately in 2011. This suggests that hospitals increased administration costs in the first year audits were expanded, perhaps reflecting large, one-time investments made to mitigate further audits. A back-of-the-envelope calculation comparing the government savings to hospital compliance costs of increasing the audit rate finds that for every \$1000 in government savings from the RAC program between 2011 and 2015, hospitals had to spend \$173 in compliance costs. A large majority of government savings from the RAC program (73 percent) are due to deterred

[3]There is evidence that hospitals substitute observation stays for hospital admissions in order to avoid a patient stay to be counted as a readmission by HRRP (Sheehy et al., 2021).

future admissions.[4]

One potential source of higher administration costs is the investments used to identify sources of unnecessary care. Figure 4d and Table 2 column 5 present coefficients for whether a hospital reported installing tracking software related to RAC audits: medical necessity checking software. Medical necessity checking software validates medical necessity by cross-referencing patient and stay details with payer rules in real-time in order to reduce payment errors and denials. Hospitals can use medical necessity checking software to identify potential vulnerabilities that RACs could later penalize. In response to a one percentage point increase in RAC audit rate in 2011, hospitals are 2.2 percentage points more likely to report that they are installing or upgrading their medical necessity software after 2012 (3.7 percent of the 2010 share of hospitals with this software installed). The results suggest that RAC auditing increased the administrative burden of hospitals, in line findings from the AHA RACTrac survey. Hospitals face a "double whammy" of both reduced inpatient revenue and higher compliance costs. While RAC audits saved money for Medicare by correcting erroneous payments for previous admissions and deterring future admissions, the burden of identifying unnecessary admissions fell onto hospitals.

Figure 4 also illustrates the dynamics of hospital responses to auditing. In response to a higher audit rate in 2011, admissions and inpatient revenue decline at first and then plateau. Even in 2014 and 2015, when RAC inpatient audit volume decreased significantly (Figure 1), admissions and revenue did not recover to their pre-auditing levels. This suggests that experiencing high initial audit activity in 2011 had longer-term effects on hospital behavior several years down the line, even absent high levels of contemporaneous audit activity. In contrast to the admissions and revenue response, the compliance cost response occurs faster, with an immediate increase in hospital administration spending in 2011. The short-term burst in compliance costs but longer-term decline in admissions is in line with hospitals making upfront, one-time investments in response to auditing. One example of such an expense would be the installation or upgrading of software, which is what we see happening in subsequent years. Because providers' compliance costs include investments, then the time horizon used in the welfare or cost-benefit analysis matters – the compliance cost of an investment should be compared to the present discounted value of any savings in future years attributed to the investment.

The burst of upfront compliance costs coupled with the slower but permanent decline in admissions imply that after audits began, hospitals went through a costly adjustment period

[4] Using the coefficients from Table 2 for a one percentage point increase in audit rate, I calculate that the present discounted value of total government savings between 2011 and 2015 for the median hospital is \$2.2m and the present discounted value of corrected prior admissions is \$597k. The present discounted value of compliance costs associated with a one percentage point increase in audit rate is \$372k.

to modify their admission practices in order to mitigate future audits. This suggests that prior to RAC audits, these hospitals were not knowingly admitting unnecessary cases – they just did not know exactly which cases would be considered necessary or not. If hospitals were committing intentional fraud (i.e., they knew which admissions were unnecessary but still admitted them nonetheless), then we would expect to see an immediate decline in unnecessary admissions in response to audits.

In Figures A10 and A11, I next explore heterogeneity of the results across different types of hospitals. For-profit hospitals reduce slightly admissions more than non-profit and government hospitals (Figures A10a and A10b), and their drop in admissions is sustained and even larger at the end of the study period in 2015. Hospitals that are part of larger hospital systems respond similarly as independent hospitals (Figures A10c and A10d). Splitting hospitals by their 2011 demand rate, or the share of audited claims that resulted in an overpayment or underpayment demand, hospitals with lower demand rates take longer to reduce admissions that hospitals with relatively higher demand rates.[5] Hospitals with higher demand rates do not immediately reduce their admissions. One possible explanation is that they need time to invest in technology that to identify unnecessary admissions – hospitals with relatively higher demand rates are also more likely to install/upgrade medical necessity software (Figure A11d) in 2011 and 2012, and their administrative spending remains at elevated levels even by the end of the period (Figure A11c). The increase in administration costs and the adoption of medical necessity software is driven mostly by the 36 percent of hospitals in the sample that do not have any medical necessity software installed in 2010 (Figures A11a and A11b).

Figures A12 and A13 plot the event studies of a robustness test in which instead of regressing on instrumented audit rate, I instrument for the denial rate, or the share of all claims that are audited and a payment is demanded from. Equation 5 defines the relationship between denial rate, audit rate, and demand rate.

$$Denial\ Rate_{ht} = \underbrace{P(Audit)}_{Audit\ Rate}{}_{ht} \times \underbrace{P(Demand|Audit)}_{Demand\ Rate}{}_{ht} \tag{5}$$

Since 41 percent of audits in 2011 resulted in a demand in the main sample, we would expect that the hospital response to a one percentage point increase in denial rate should be about twice the response to one percentage point increase in audit rate. Indeed, this is what we see – for example, hospitals reduced admissions by 2.5 percent in 2012 in response to a 1pp increase in 2011 audit rate, and reduced by 5.7 percent in 2012 in response to a 1pp increase in denial rate. The denial rate results track with audit rate, and combined with

[5] The average 2011 demand rate for hospitals in the main sample was 41 percent.

the heterogeneity by demand rate results, demonstrate that hospitals are mostly responding to *audit* rate rather than *demand* rate. Hospitals with relatively low demand rates still reduce admissions in response to audits, even though a lower share of their audited claims are denied. Hospitals with low demand rates may still reduce admissions because simply going through the auditing process is costly – even if a claim is not eventually denied. So as hospitals learn about what RACs are targeting, they reduce admissions to deter future audits, not just future denials. Hospitals may want to avoid audits if the audit process itself is costly, which is consistent with other work that has shown that the "back-and-forth" interactions between providers and payers in the claim denial process is costly for providers, even when it doesn't result in a denial (Dunn et al., 2021).

4.2 Within-Hospital Results

After establishing the hospital-level effects of auditing on admissions, revenue, costs, and IT adoption, I next turn to a within-hospital, patient-level analysis of RAC audits on admission likelihood and health outcomes. To study the effect of auditing on patient outcomes, I switch to a patient-level empirical strategy. I do this because if hospitals change the number of patients they admit, then patient composition will mechanically change. Comparing patient outcomes across hospitals risks misattributing patient composition effects as the causal effect of monitoring on a patient health. With the within-hospital strategy, I compare patients within the same hospital who, depending on when they arrive at the ED, are more or less likely to be affected by the "Two Midnights Rule."

No-Manipulation Assumption A key assumption of the within-hospital approach is that hospitals do not manipulate ED arrival hour in response to the Two Midnights rule. This assumption would be violated if, for example, hospitals misreported after-midnight ED arrivals as before-midnight ED arrivals. This would invalidate a comparison of before- and after-midnight ED arrivals pre- and post-policy, as the composition of the two patient groups would change as a result of the Two Midnights rule. I argue that practically speaking, it is difficult for hospitals to manipulate ED arrival hour.

I can also test for evidence of manipulation empirically by looking at how the share of ED patients arriving before- and after-midnight changes when the Two Midnights rule is implemented. It should be easier for hospitals to manipulate ED arrival hour for arrivals closer to midnight than for arrivals further from midnight. So if hospitals were manipulating ED arrival hour to game the Two Midnights rule, we should see a larger drop in the after-midnight share among patients who arrive within an hour of midnight, relative to the after-midnight share among patients who arrive further away from midnight. Define $(Within\ 1\ Hr\ MN)_{h(i)}$ to be a dummy equal to one for patients who arrive within an hour of midnight (11PM-1AM) and zero for patients who arrive four hours away from midnight (either 8PM-9PM or 3AM-4AM). If hospitals manipulated ED arrival hour, then we would expect the coefficient on $(Within\ 1\ Hr\ MN)_{h(i)} \times (Post\ Aug\ 2013)_{t(i)}$ in Equation 6 to be negative. Table B5 reports the result of this regression – the coefficient is insignificant and positive. This suggests that hospitals did not manipulate ED arrival hour in response to the Two Midnights rule, supporting the validity of the within-hospital approach.

$$
\begin{aligned}
After\ MN_{a(i)} = \delta_0 + \delta_1 (Within\ 1\ Hr\ MN)_{h(i)} (Post\ Aug\ 2013)_{t(i)} \\
+ Hosp_{h(i)} + Qtr_{t(i)} + Hosp_{h(i)} \times Qtr_{t(i)} + \varepsilon_i
\end{aligned} \tag{6}
$$

P(Inpatient) After substantiating the no-manipulation assumption, I proceed with the difference-in-difference analysis on admission likelihood. I first plot the average share of patients admitted from the ED by whether the patient arrived before or after midnight in Figure A14. Before the Two Midnights rule goes into effect in 2013Q3, after-midnight arrivals are on average more likely to be admitted than before-midnight arrivals. This gap closes abruptly as soon as the Two Midnights rule goes into effect in 2013Q3. The dynamic difference-in-difference results from estimating Equation 4 reflect this pattern as well (Figure 5). Table 4 reports the aggregate $(Post\ Aug\ 2013)_{t(i)} \times (After\ MN)_{h(i)}$ coefficients. Note that there is no clear trend in the pre-policy (before 2013Q3) coefficients. This which supports the parallel trends assumption that absent the Two Midnights rule, these outcomes would have trended similarly. Panels 5a and 5b show that immediately after the Two Midnights rule goes into effect in 2013Q3, there is an immediate drop in the relative share of after-midnight Medicare ED patients who are admitted as inpatient, and a symmetric sharp increase in the share placed into observation (and never admitted). In columns 1 and 2 of Table 4, the coefficients on the inpatient dummy and observation dummy 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 status. Panel 5c and column 3 in Table 4 show that there was no change in the percent of patients being sent home directly from the ED ("Out"). For ED patients on the margin of being admitted inpatient, hospitals still preferred to keep them in the hospital rather than sending them home directly. Column 5 of Table 4 considers whether there is an effect on non-Medicare patients and finds that non-Medicare patients do not face a reduction in admissions because of the Two Midnights rule. This indicates that there were no spillovers from Two Midnights rule onto populations not covered by the rule.

Similarly to the across-hospital approach, we find that hospitals respond to increased audit rates at the patient level by reducing inpatient admissions. But the patient-level results do differ from the hospital-level ones in that at the patient level, we find substitution from inpatient admissions to observation stays. The difference may be driven by the subset of patients I focus on in the within-hospital strategy: ED patients who have already arrived at the hospital. Conditional on arriving at the hospital, providers may be hesitant to send patients home, and so they place them into observation stays. But the across-hospital finding of a decrease in inpatient admissions without an increase in outpatient revenue could be the result of hospital-level efforts to reduce inpatient admissions *before* patients even arrive at the hospital, like discouraging physician referrals and transfers or choosing not to expand their emergency department.

P(Revisit) Next, I consider whether the reduction in inpatient admissions negatively affected patient health outcomes. Patients who arrived to the ED after midnight were less likely to be admitted because they were more likely to be audited – what was the effect of being denied admission on their health outcomes? One measurable proxy for patient health after an ED visit is if the patient revisits a hospital shortly after, either through the ED or inpatient.[6] Panel 5d plots the results for a dummy indicating whether a patient revisited a hospital within 30 days of her ED visit, and column 4 in Table 4 reports the aggregated coefficient. After-midnight patients were not more likely to revisit a hospital within 30 days after the Two Midnights rule came into effect, despite their reduced admission likelihood. Because we've shown that hospitals are not manipulating ED arrival hour, that this result is not driven by a change in after-midnight patient composition post-policy.

Heterogeneity by Patient Severity The average results suggest that on average, the reduction in admissions does not correspond to an increased likelihood of a revisit within 30 days. However, a null average effect on patient outcomes may mask heterogeneity, as only a subset of patients are on the margin between inpatient admission and not. There is substantial heterogeneity in admission rates by patient severity – ED patients whose cases are very severe will always be admitted, while ED patients whose cases are very mild will never be admitted. Neither severe nor mild patients will be on the margin when it comes to a policy aimed at reducing unnecessary admissions. Instead, it is patients in the middle of the severity distribution who are most likely to be denied admission because of policies like RAC audits, so we would also expect that any negative effects on patient outcomes will be concentrated among these patients as well. Focusing on the heterogeneous effect for marginal patients may reveal effects on patient outcomes that otherwise may be obscured in

[6]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.

the average effect.

So, I next explore heterogeneity by patient severity. I predict a patient's severity based on information available to providers before an inpatient admission. Using data on ED visits between 9AM and 3PM, outside of the time window for the main results, I estimate a logistic regression predicting whether a patient is admitted within 30 days of an ED visit based on patient demographics, hospital and quarter, and information on current and previous hospital visits. Patient demographics include age-bin, sex, race, Hispanic dummy, point of origin dummy, and mean zip code income. Information on the 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. I then apply the coefficients from this estimation to the main sample to create a measure of predicted patient severity, and split patients into deciles of predicted severity. I then re-estimate the specification in Equation 3, interacting $(After\ MN)_{a(i)} \times (Post\ Aug\ 2013)_{t(i)}$ with a dummy for each predicted risk decile.

Figure 6 plots the heterogeneity results for inpatient status and revisits within 30 days and the coefficients are reported in Table B6. Inpatient admission likelihoods are unaffected by RAC audits for patients with the lowest and the highest risk deciles, and the reduction in admissions is concentrated among patients with risk deciles in the middle of the severity distribution. In contrast, we do not see this pattern for revisits within 30 days – the effect on revisits is statistically insignificant at all risk deciles. Thus, the overall null effect on revisits is not masking heterogeneity by patient severity, and health outcomes are unaffected even for patients most likely to be denied admission due to the Two Midnights rule.

In sum, using the within-hospital approach, I find that at the patient-level, a higher audit threat reduces inpatient admissions but does not affect patient outcomes. These results support the notion that auditing has no first-order effects on patient outcomes. In the context of the ED, where hospitals are obligated to treat every patient who walks through the door, hospitals substituted inpatient admissions with observation status. They do so in a targeted way, denying admission only for patients in the middle of the severity distribution. By switching from inpatient admission to outpatient observation, hospitals reduced their Medicare revenue as on average observation stays are much less expensive than inpatient stays (Medicare Payment Advisory Commission, 2015). The switch in patient status did not seem to negatively affect patient health outcomes. Hospitals reoptimized in response to the threat of audit in a way which reduced Medicare spending without affecting patient outcomes. This indicates that most of the welfare effect of the RAC audits is through net savings to the government and provider compliance costs, and not through direct effects on patient health outcomes.

5 Marginal Welfare Analysis

The empirical results demonstrate that in response to higher audit rates, hospitals reduce admissions and increase compliance costs, with no evidence of patient harm. But with just the empirical estimates alone, it is unclear what the net welfare effect of a higher audit rate would be, as a higher audit rate is associated with both government savings and private compliance costs. I next bring these estimates together to estimate the welfare effect of a marginal increase in audit rate at the time audits were introduced in 2011.

I use a sufficient statistics-style framework that adapts the model of tax administration in Keen and Slemrod (2017) to the Medicare inpatient prospective payment system and RAC auditing context. The framework highlights the estimates required to determine the marginal welfare effect of introducing a higher audit rate. It is important to note that since the empirical results are estimated from a major expansion of auditing scope in 2011, these marginal welfare effects should be interpreted as answering the question, "What if RACs had increased audit rates *in 2011* when they expanded the audit scope?" rather than interpreting them as the welfare estimates of a marginal increase in audit rate *today*.

5.1 Framework

I first present the hospitals' objective function. In contrast to profit-maximizing firms, I assume that hospitals are altruistic in that they care about patient benefit as well as revenue from treating patients, net of costs (Chang and Jacobson, 2012). When audits began, RACs could audit prior admissions in the last three years but hospitals could only change future admissions. To capture this, I split admissions into prior admissions n_P, which the hospital admitted before audits were introduced, and future admissions n_F, which the hospital decides after auditing begins. In total, $n_P + n_F$ admissions are at risk of audit. The government audits share a of $n_P + n_F$ admissions. Net revenue is comprised of the revenue from treating n_F future patients less the amount reclaimed from audits n_P previous claims $R(n_P + n_F, a)$, the compliance costs of audits $k(a, n_P + n_F)$, and the cost to treat n_F future patients $c(n_F)$. The patient benefit from future admissions is $b(n_F)$.

Hospital's objective function:

$$\max_{n_F} \Pi \left(\underbrace{R(n_P + n_F, a) - k(a, n_P + n_F) - c(n_F)}_{\text{net hospital revenue}}, \underbrace{b(n_F)}_{\text{patient benefit}} \right) \tag{7}$$

To determine the welfare effect of introducing a higher audit rate, I next lay out the social welfare function. For simplicity, I assume that the social welfare function is additively separable in its three components: the (1) hospital's objective function; (2) the societal value $V(\cdot)$ of government revenue net of spending on inpatient stays and administrative costs; and (3) the societal value $\gamma(\cdot)$ of the patient benefit of future admissions.

Social welfare function:

$$W = \underbrace{\left\{\max_{n_F} \Pi(n_P, a)\right\}}_{\text{hospital objective fxn}} + \underbrace{V\left(G - R(n_P + n_F, a) - m(a, n_P + n_F)\right)}_{\text{net government revenue}} + \underbrace{\gamma\left(b(n_F)\right)}_{\text{patient benefit}} \tag{8}$$

Taking the derivative of the social welfare function with respect to audit rate a and applying the envelope theorem, we get the following:

$$\underbrace{(V' - \Pi')}_{\substack{\text{marginal value of public} \\ \text{funds vs. marginal value} \\ \text{of hospital revenue}}} \times \underbrace{(-R_a)}_{\text{marginal revenue}} = \underbrace{k_a}_{\substack{\text{marginal hospital} \\ \text{compliance cost}}} + \underbrace{V' m_a}_{\substack{\text{weighted marginal} \\ \text{gov't admin cost}}} + \underbrace{\gamma' \frac{db}{dn_F} \frac{dn_F}{da}}_{\substack{\text{weighted value of} \\ \text{patient benefit}}} \tag{9}$$

where the left-hand side of Equation 9 represents the societal value of government savings from auditing and the right-hand side represents the societal costs. Audits facilitate a transfer from hospitals back to the government, and this transfer is only valuable if the marginal value of public funds is greater than the marginal value of hospital revenue ($V' > \Pi'$). The right-hand side of Equation 9 captures the costs of a higher audit rate through hospital compliance costs, government administrative costs, and patient health effects. If the left-hand side is less than the right-hand side, then a higher audit rate is welfare-decreasing. Vice versa, if the right-hand side is less than the left-hand side, then a higher audit is welfare-increasing. Equation 9 shows that to calculate the marginal welfare effect of increasing auditing, we need four key estimates: the marginal effect on government revenue R_a, the marginal hospital compliance cost k_a, the marginal government administrative cost of conducting audits m_a, and the marginal effect on patient benefit $\frac{db}{dn_F} \frac{dn_F}{da}$.

The time horizon that we are considering is important as well – namely, how many years of savings and costs should be included in the calculation? If hospitals incur fixed costs like a large upfront investment that lasts for several years, then we should compare it to the (present discounted value) of savings over a multi-year horizon. To remain agnostic about the time horizon for calculating the welfare effect, I instead ask the question: after how many years (if any) would the introduction of a higher audit rate in 2011 be welfare-improving? I therefore calculate the marginal welfare effect per hospital in each year in different years.

5.2 Calculation

In order to calculate R_a, k_a, and m_a in each year, I use estimates derived from the dynamic difference-in-difference estimates in Table 2. Section B describes this calculation in further detail. For the marginal patient benefit $\frac{db}{dn_F}\frac{dn_F}{da}$, I assume in the baseline specification that it is being denied admission to be 0. This is motivated by the null result from the patient-level empirical strategy, which is also in line with other work that finds that the marginal hospitalization has no effect on patient health (Currie and Slusky, 2020). I then relax the assumption of a null patient health effect by exploring how the marginal welfare effect varies with different marginal effects on patients and marginal values of public funds in Figure 8. Table 5 lists the parameters and estimates used to calculate the welfare effect.

Figure 7 plots $(-R_{a_T} - k_{a_T} - m_{a_T})$, or the difference between the marginal savings and marginal costs at year T from a one percentage point higher 2011 audit rate. A higher 2011 audit rate is welfare-improving if this value is positive and welfare-reducing if this value is negative. Figure 7 plots this value in three cases that decompose the overall welfare calculation: (1) audits reduce future inpatient revenue and increase compliance costs (baseline estimate); (2) no compliance costs; and (3) no reduction in future admissions.

Case (1) is the main welfare calculation – increasing auditing only became welfare-improving five years after 2011. In 2015, the savings to the government finally outweighed the costs associated with auditing; these savings were mostly driven by the reductions in future admissions rather than reclaimed payments. The estimates imply that a one percent higher 2011 audit rate results in a marginal welfare improvement of \$61,000 by 2015; across all 2,901 eligible hospitals in the U.S., this is equivalent to a welfare improvement of \$177 million. Comparing cases (1) and (2), we see the important role compliance costs play in determining the magnitude of the marginal welfare effect. Absent compliance costs, a higher audit rate is always welfare-improving. The marginal welfare effect by 2015 would be almost *7.1 times larger* if hospitals did not face any compliance costs – absent compliance costs, the marginal welfare effect by 2015 across all hospitals of a one-percent-higher audit rate would be \$1.26 billion. The gap between the welfare effects in cases (1) and (2) diminishes over time as more savings accrue through deterred admissions. Comparing case (1) to case (3), we see that the key to the positive welfare effect in the main estimate is the deterrence of future admissions. If audits simply collect money back from prior admissions but still impose the same compliance costs, a higher audit rate is always welfare-reducing, since the reclaimed payments would not cover the compliance costs.

In Figure A15 I compare the "most conservative" case to the "least conservative" one, where the most conservative case would correspond to the highest costs and lowest savings. In the most conservative case, RACs charge the highest contingency fee of 12.5 percent, the deterrence effect on admissions is 0 after 2015, and CMS has to refund 68 percent of demanded payments. In the least conservative case, RACs demand the lowest contingency fee of 9 percent, the effect on admissions is permanently negative after 2015, and CMS keeps all the demanded payments. Even in the most conservative case, increasing audits is welfare-improving by 2015. The parameters used for this robustness test are reported in Table B7.

These welfare results hinge on two assumptions: first, that the marginal effect on patient health of being denied admission is zero, and second, that the marginal value of public funds is 1.3, relative to a marginal value of hospital revenue of 1. To explore how these assumptions affect the findings, Figure 8 plots the relationship between marginal welfare per hospital in 2015, the marginal effect on patient health, and the marginal value of public funds. At a marginal value of public funds of 1.3, increasing the audit rate is still welfare-improving by 2015 as long as the harm per patient denied admission is less than $245. Figure A16 plots this relationship in 2011 and 2018.

In interpreting these welfare estimates, it is important to acknowledge that RAC auditing program's implementation was atypical. The RAC program was controversial and its rollout was mired with multiple pauses, policy changes, and lawsuits. Thus the estimated effects and implied welfare analysis should be viewed through this lens – the findings reflect hospitals' responses to an imperfect implementation of the RAC program. RAC audits of inpatient claims were paused in 2014 and declined after, never recovering to their 2012 peak levels.[7] Section A.3 covers the timeline of the RAC program in more detail, but in short CMS ran into many unexpected roadblocks in implementing RAC audits.

6 Conclusion

Governments often monitor the third parties they contract with to ensure compliance and cost-effectiveness of public expenditure. The welfare effect of increased monitoring depends on the money it saves, the costs to run or comply with monitoring, as well the spillovers it has on other parties. I study these outcomes in the context of Medicare audits of hospital stays, combining data on audits, hospital behavior, and patient-level outcomes with two identification strategies, one at the hospital level and one at the patient level. I find that monitoring induces hospitals to reduce expenditure through fewer admissions and show evidence that hospitals targeted their response to only cut back on unnecessary admissions.

[7]According CMS at the time, the pause was temporary and audits could resume at any point. 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 provider burden.

At the hospital level, I find that hospitals subject to exogenously higher audit rates reduce admissions, and these reductions persist for many years. The reductions in admissions were driven by short stays and stays with circulatory diagnoses, which CMS identified as being more likely to be unnecessary. At the patient level, hospitals are less likely to admit patients who, if admitted, have a greater probability of being audited. But despite being denied admission, these patients were not more likely to revisit the hospital at a later date, suggesting that patients denied admission because of audits did not face worsened health outcomes.

While I do not find evidence that the reduction in admissions caused patients enough harm to warrant a revisit to the hospital, it did come at a substantial private compliance cost to hospitals. In response to increased monitoring, hospitals increased administration spending as they invested in technology to detect and reduce unnecessary admissions. A back-of-the-envelope calculation finds that for every \$1000 in Medicare savings over five years, hospitals spent \$173 in compliance costs. The savings from monitoring accrued over several years, mostly driven by sustained reductions in unnecessary admissions. Given the high up-front compliance costs for providers and the fact that the savings to Medicare accrued over time, monitoring through the RAC program is welfare-improving only after five years.

The findings highlight an unintended consequence of healthcare policy: the potential for well-intentioned policy to increase provider compliance costs. Reducing unnecessary government expenditure without harming patients is not sufficient for a policy to be welfare-improving – it must also not be extremely costly to implement the policy or for third parties to comply with the policy. In the case of the RAC program, the welfare gain from increasing monitoring is much smaller once we take into account hospitals' compliance costs – the marginal welfare effect of increasing audits would be almost 7.1 times larger if hospitals did not incur any compliance costs. This is especially pertinent in the U.S., where healthcare providers' paperwork burden is already relatively high (Cutler and Ly, 2011; Himmelstein et al., 2014; Papanicolas et al., 2018). In this case, I document an instance where the third parties incurred private costs to *save* money on behalf of the government. Overall, these findings highlight the importance of accounting for all sources of savings and costs, both public and private, in evaluating policy.

References

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.

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.

Arrow, Kenneth J., "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, 1963, *53* (5), 941–973. Publisher: American Economic Association.

Berwick, Donald M. and Andrew D. Hackbarth, "Eliminating waste in US health care," *JAMA*, April 2012, *307* (14), 1513–1516.

Besley, Timothy and Stephen Coate, "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs," *The American Economic Review*, 1992, *82* (1), 249–261. Publisher: American Economic Association.

Boccuti, Cristina and Giselle Casillas, "Aiming for Fewer Hospital U-turns: The Medicare Hospital Readmission Reduction Program," March 2017.

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

Centers for Medicare and Medicaid Services, "CMS Glossary," May 2006.

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

_ , "Guidance on Hospital Inpatient Admission Decisions," MLN Matters SE1037 July 2012.

_ , "2018 Medicare Fee-for- Service Supplemental Improper Payment Data," Technical Report 2019.

_ , "Maryland All-Payer Model," 2021.

Chang, Tom and Mireille Jacobson, "What do Nonprofit Hospitals Maximize? Evidence from California's Seismic Retrofit Mandate," in "in" August 2012, p. 61.

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.

Currie, Janet, "The Take-up of Social Benefits," in "Poverty, the Distribution of Income, and Public Policy, (New York: Russell Sage)" 2006, pp. 80–148.

_ **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.

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.

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.

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.

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.

Finkelstein, Amy, 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.

Gross, Tal, Adam Sacarny, Maggie Shi, and David Silver, "Regulated Revenues and Firm Behavior: Evidence from a Medicare Overhaul," in "in" April 2021.

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

Himmelstein, David U., Miraya Jun, Reinhard Busse, Karine Chevreul, Alexander Geissler, Patrick Jeurissen, Sarah Thomson, Marie-Amelie Vinet, and Steffie Woolhandler, "A Comparison Of Hospital Administrative Costs In Eight Nations: US Costs Exceed All Others By Far," *Health Affairs*, September 2014, *33* (9), 1586–1594. Publisher: Health Affairs.

Holmström, Bengt, "Moral Hazard and Observability," *The Bell Journal of Economics*, 1979, *10* (1), 74–91. Publisher: [RAND Corporation, Wiley].

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

Jaffe, Susan, "FAQ: Hospital Observation Care Can Be Costly For Medicare Patients," August 2016.

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.

Keen, Michael and Joel Slemrod, "Optimal tax administration," *Journal of Public Economics*, August 2017, *152*, 133–142.

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.

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.

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

_ , "Hospital Inpatient and Outpatient Services," Technical Report March 2019.

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

Miller, Mark, "Hospital Short-Stay Policy Issues," May 2015.

Papanicolas, Irene, Liana R. Woskie, and Ashish K. Jha, "Health Care Spending in the United States and Other High-Income Countries," *JAMA*, March 2018, *319* (10), 1024–1039.

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.

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., Charles Locke, Jeannine Z. Engel, Daniel J. Weissburg, Stephanie Mackowiak, Bartho Caponi, Sreedevi Gangireddy, and Amy Deutschendorf, "Recovery Audit Contractor Audits and Appeals at Three Academic Medical Centers," *Journal of Hospital Medicine*, April 2015, *10* (4), 212–219.

_ , Farah Kaiksow, W. Ryan Powell, Andrea Gilmore Bykovskyi, Christie M. Bartels, Blair Golden, and Amy Jh Kind, "The Hospital Readmissions Reduction Program and Observation Hospitalizations," *Journal of Hospital Medicine*, July 2021, *16* (7), 409–411.

_ , Jeannine Z. Engel, Charles F. S. Locke, Daniel J. Weissburg, Kevin Eldridge, Bartho Caponi, and Amy Deutschendorf, "Hospitalizations With Observation Services and the Medicare Part A Complex Appeals Process at Three Academic Medical Centers," *Journal of Hospital Medicine*, April 2017, *12* (4), 251–255. WOS:000399578000007.

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.

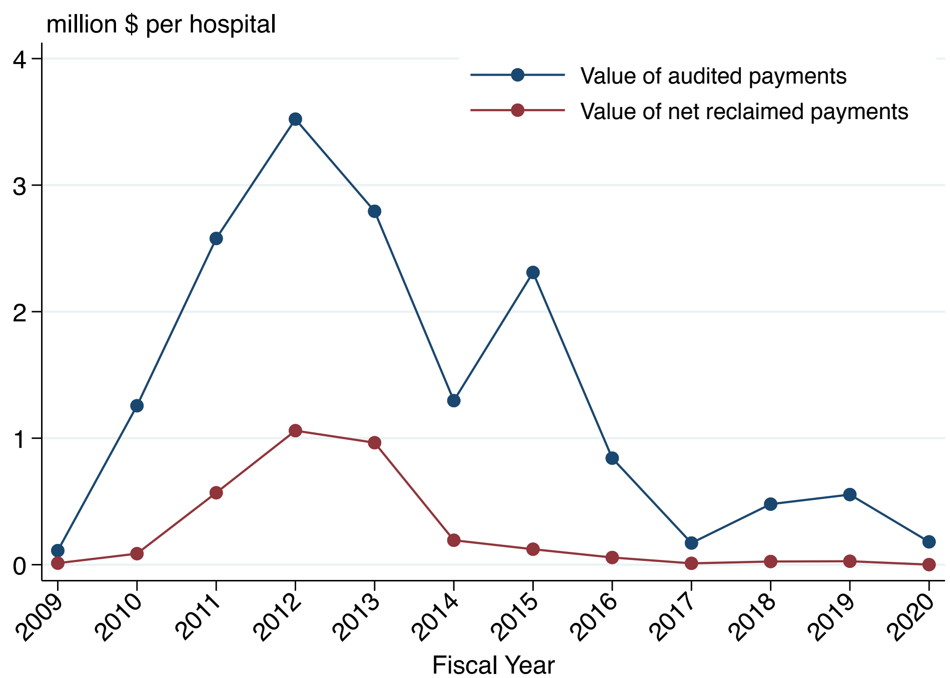
U.S. Department of Health and Human Services, "Maryland All-Payer Model," 2019.

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.

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

7 Figures

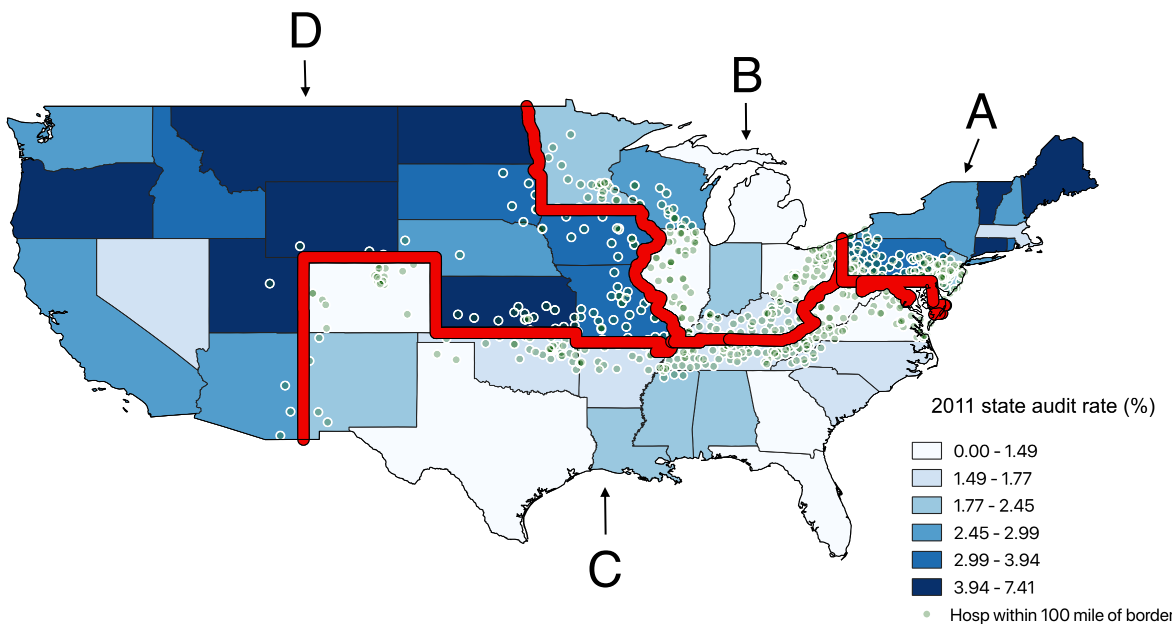
Figure 1. Value of Audited Inpatient Payments and Net Reclaimed Payments per Hospital, by Year of Audit



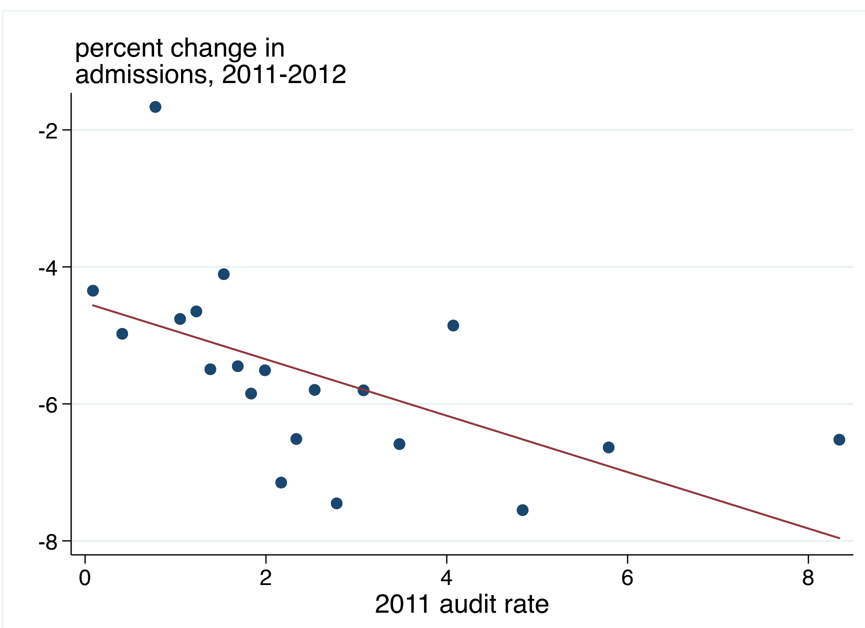
This figure plots the average per-hospital value of inpatient payments audited by RACs and the net reclaimed payments, by year of audit. Net reclaimed payments are defined as the sum of reclaimed payments from overpayments minus refunded payments from underpayments. These values are based on RACs' original reclaimed or refunded payments at the time of audit, and do not take into account the 2014 settlement CMS reached with some hospitals to partially refund denied claims in exchange for dropping appeals.

Figure 2. Audit Rates

(a) Average 2011 audit rates by state



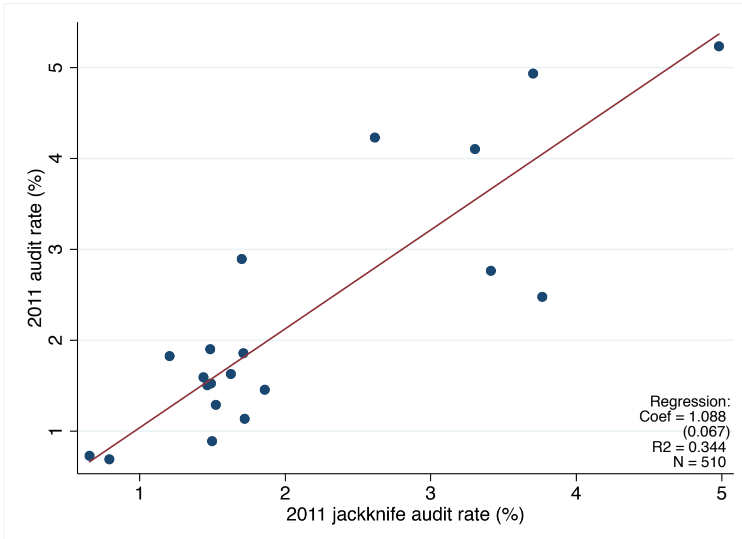
(b) Relationship between 2011 hospital audit rate and percent change in Medicare admissions, 2011-2012



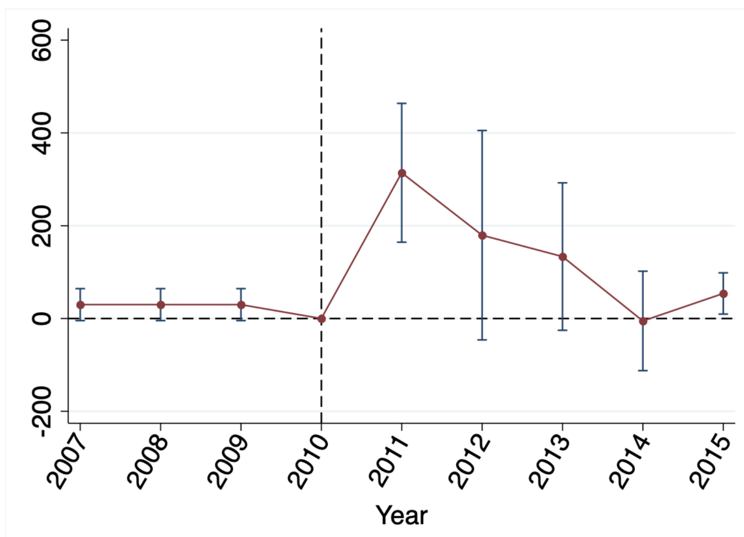
The top panel plots the 2011 average state audit rates, where audit rate is defined as the percent of a hospital's 2008-2011 claims that were audited by RACs. The RAC regions are: Region A (Northeast), Region B (Midwest), Region C (South), and Region D (West). Darker shades denote higher audit rate. The red line demarcates RAC regions. The green dots symbolize hospitals within 100 miles of RAC border. Maryland was not audited under the RAC program as it uses a unique all-payer rate-setting system for hospital services (Centers for Medicare and Medicaid Services, 2021). The bottom panel plots the binscatter of the relationship between 2011 hospital audit rate and the percent change in Medicare admissions between 2011 and 2012.

Figure 3. First Stage

(a) 2011 jackknife audit rate and 2011 audit rate



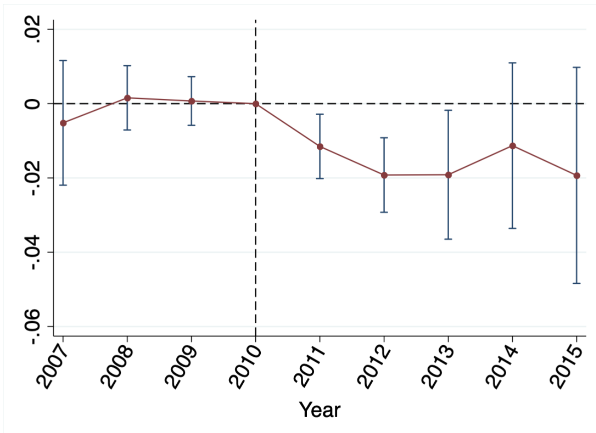
(b) Effect of instrumented 2011 audit rate on amount demanded ($)



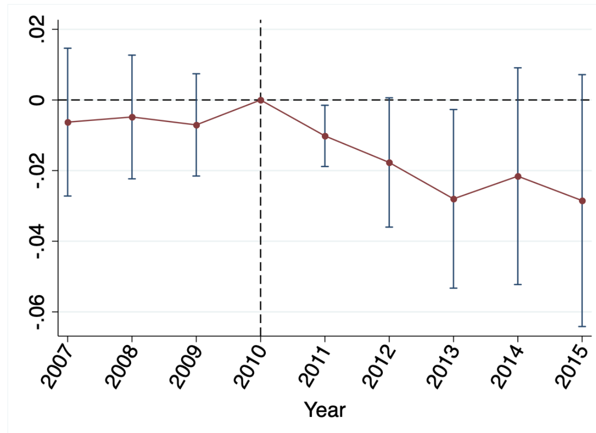
The top panel of this figure plots a binscatter of 2011 actual hospital audit rate compared to 2011 jackknife state audit rate. 2011 audit rate is defined as the share of 2008-2011 inpatient claims that were audited by RACs in 2011. Jackknife state audit rate is defined as the average audit rate of all other hospitals in the same state as a given hospital. The bottom panel of this figure plots the event study of the IV coefficient and 95% confidence interval of the specification in Equation 2, where the outcome variable is the amount demanded ($) from audits of inpatient claims per hospital. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group."

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

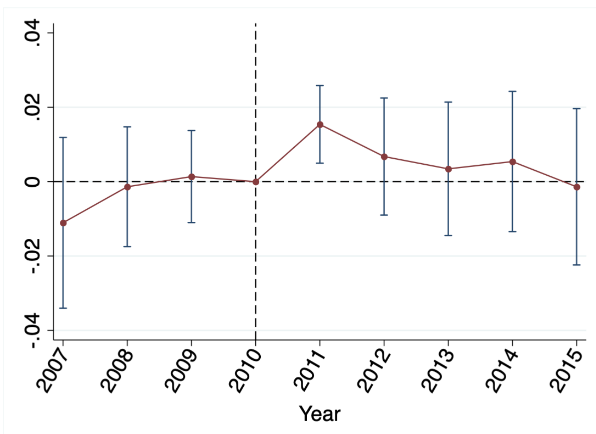
(a) Log Medicare admissions



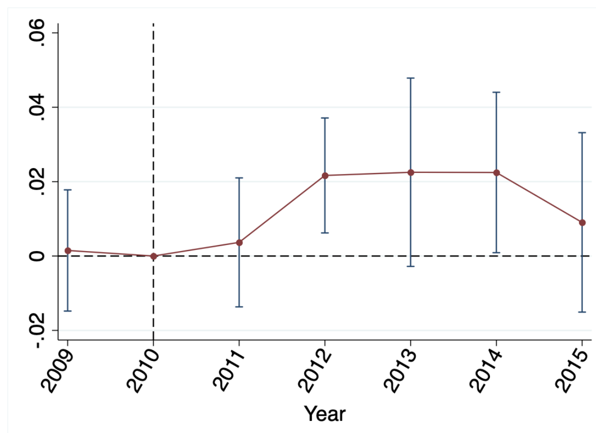
(b) Log Medicare inpatient revenue



(c) Log hospital administrative costs

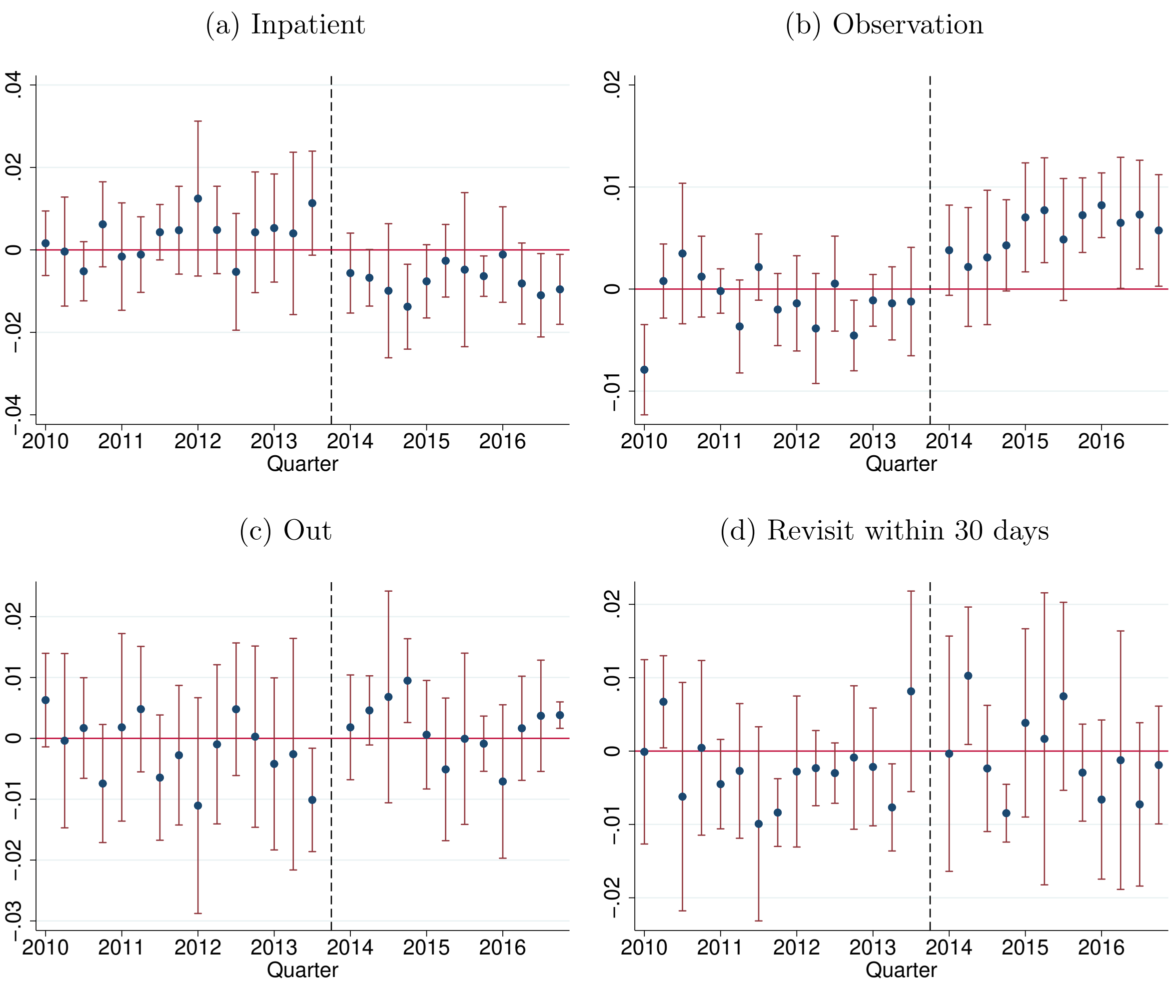


(d) Dummy for installing medical necessity software



This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2. The omitted year is 2010. Each coefficient estimates the effect of a 1pp 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 administration costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Administration cost share is defined as net administrative costs divided by net total costs for a hospital. Dummy 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group."

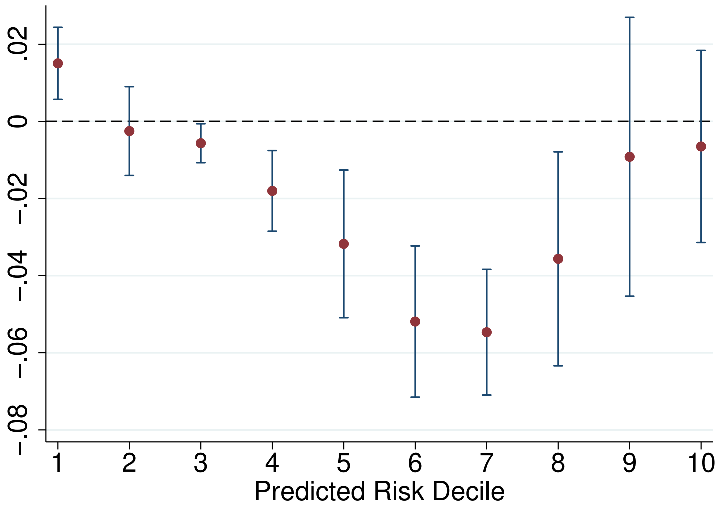
Figure 5. Event Studies on After-Midnight ED Arrival Dummy on Patient Status and Outcomes



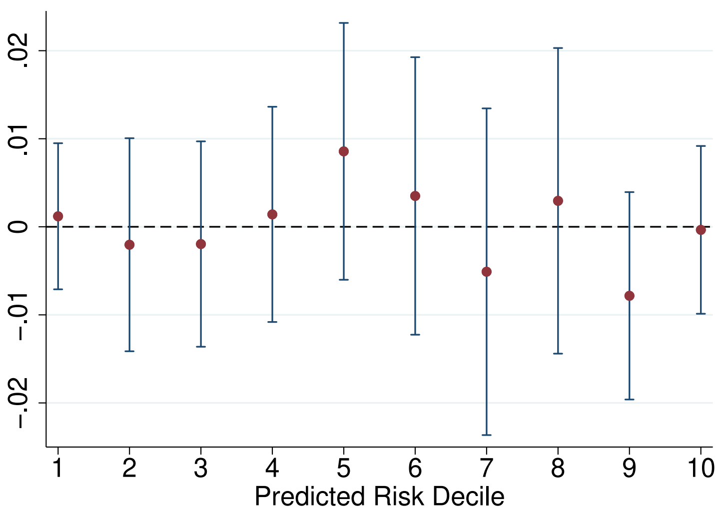
This figure plots the quarterly difference-in-difference coefficients of $AfterMN_{a(i)} \times Qtr_{t(i)}$ and 95% confidence intervals of the specification in Equation 3. The results are clustered at the ED arrival hour and quarter level. $AfterMN_{a(i)}$ is a dummy if a visit occurred after midnight (midnight-3AM). The omitted quarter is 2013Q3. "Admitted" is a dummy for whether the patient was eventually admitted as inpatient from the ED. "Observation" is a dummy for whether the patient was placed in observation status and was never admitted. "Out" is a dummy equal to one when a patient is neither admitted nor placed in observation status. "Revisit within 30 days" is a dummy 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 dummy, point of origin dummy, last ED visit within 30 days dummy, number of chronic conditions, and zip code income.

Figure 6. Heterogeneity of Two Midnights Rule Effect by Patient Severity

(a) Inpatient

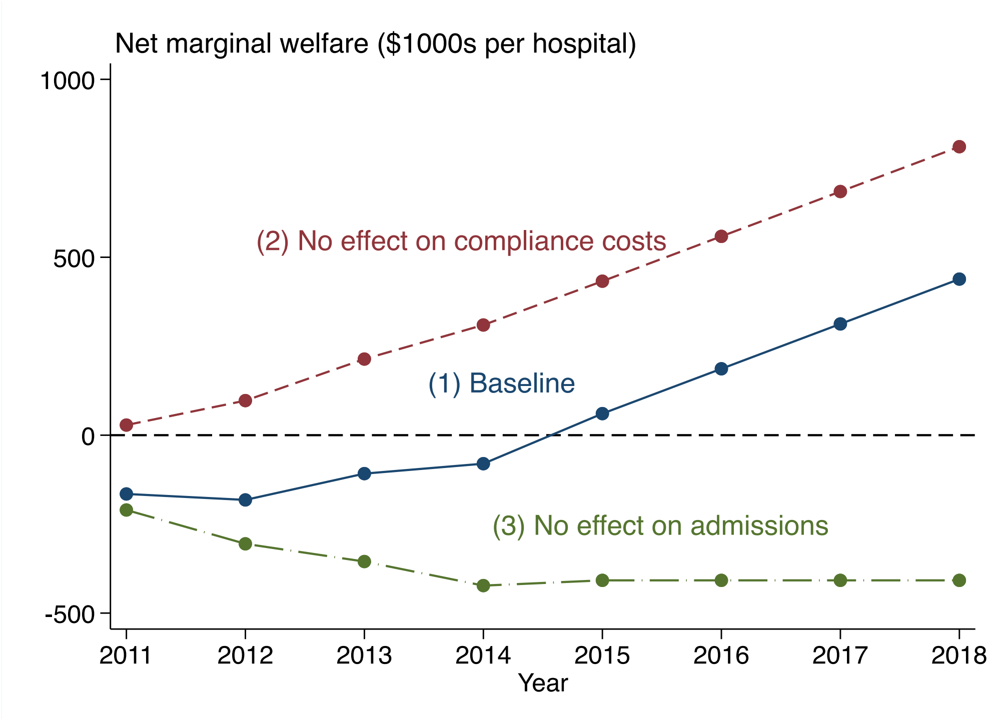


(b) Revisit in 30 days



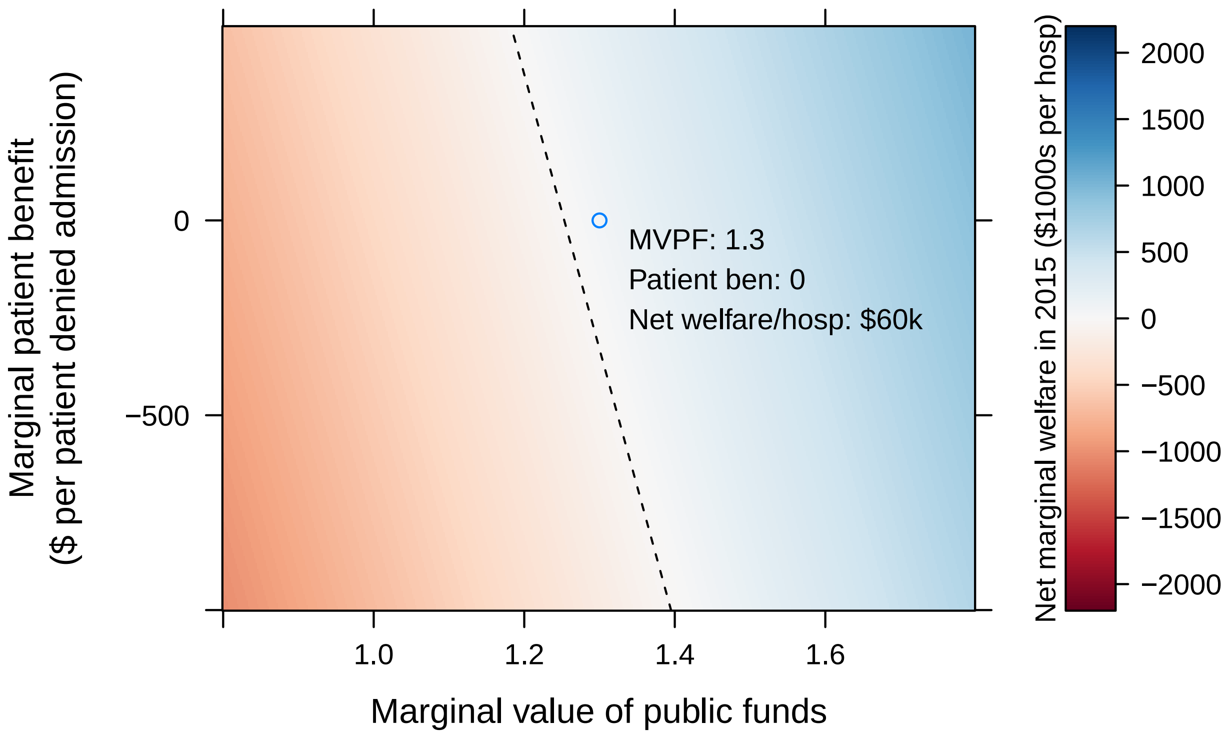
This figure plots estimates and 95% confidence intervals of the $Post_{t(i)} \times AfterMN_{a(i)}$ coefficient in Equation 3, interacted with a dummy for predicted severity decile. The top panel plots results for a dummy for whether the patient was admitted as inpatient from the ED, and the bottom panel plots results for a dummy for whether the patient revisited any hospital in Florida in 30 days after 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 9AM and 3PM of a dummy 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 dummy, point of origin dummy, 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 7. Welfare Analysis Estimates



This figure plots the per-hospital welfare effect, or the difference between the savings and compliance/administrative costs of auditing, over time of increasing 2011 audits. Increasing audits is welfare-improving if this value is positive and welfare-reducing if this value is negative. This figure plots this value in three cases: (1) audits reduce future inpatient revenue (permanently) and increase compliance costs (the main estimate); (2) audits reduce future inpatient revenue with no compliance costs; and (3) audits do not reduce future inpatient revenue and have no effect on compliance costs. Table 5 lists the parameters and estimates used to calculate the welfare effects for each case.

Figure 8. Marginal Welfare Effect in 2015 by Patient Benefit and Marginal Value of Public Funds



This figure plots the per-hospital marginal welfare effect of increasing 2011 audits, with varying assumptions about the marginal value of public funds and the marginal patient benefit ($ per patient denied admission) in 2015. Increasing audits is welfare-improving if this value is positive (blue) and welfare-reducing if this value is negative (red). Table 5 lists the other parameters and estimates used to calculate the welfare effects. The blue point represents the baseline specification, which assumes a MVPF of 1.3 and no patient health effects from reduced admissions. The dotted line denotes the set of combinations of marginal patient benefit and marginal value of public funds where the marginal welfare effect is 0.

8 Tables

Table 1. Summary statistics of 2010 hospital characteristics by 2011 hospital audit rate, overall sample vs. border sample

	(1)		(2)		(3)		(4)	
	Overall Hospitals				Border Hospitals			
	Above Median		Below Median		Above Median		Below Median	
A. RAC Program Characteristics								
2011 audit rate	3.60	(1.89)	0.73	(0.65)	3.62	(2.09)	0.84	(0.63)
share in Region A	0.23		0.11		0.15		0.02	
share in Region B	0.19		0.20		0.31		0.42	
share in Region C	0.28		0.54		0.27		0.48	
share in Region D	0.30		0.16		0.28		0.09	
B. Overall Characteristics								
beds	182.04	(164.09)	228.76	(195.51)	176.66	(194.80)	181.73	(149.42)
share urban	0.68		0.76		0.49		0.60	
share non-profit	0.68		0.58		0.72		0.66	
share for-profit	0.12		0.25		0.12		0.21	
share government	0.20		0.16		0.16		0.14	
share independent	0.42		0.31		0.44		0.35	
total cost (million $)	193.78	(248.46)	215.21	(269.58)	164.60	(294.09)	163.17	(222.23)
net admin costs (million $)	28.84	(39.11)	32.12	(39.59)	24.83	(44.68)	26.18	(47.14)
share with medical necessity app.	0.67		0.68		0.73		0.68	
C. Medicare Inpatient Admission Characteristics								
admissions	3056.70	(3057.97)	3931.26	(3351.82)	3007.99	(3332.92)	3225.81	(2833.25)
mean payment ($)	8788.95	(3134.69)	9001.31	(3104.10)	7539.54	(2268.09)	7618.78	(2231.30)
total payments (million $)	30.28	(38.07)	39.03	(42.22)	26.66	(40.01)	28.13	(31.07)
mean share stays, LOS = 0-2	0.31	(0.07)	0.30	(0.07)	0.31	(0.07)	0.31	(0.06)
mean share stays, circulatory diagnosis	0.19	(0.06)	0.21	(0.08)	0.20	(0.06)	0.20	(0.05)
N neighboring hospitals					16.29	(11.29)	17.13	(11.21)
Observations	1474		1430		255		255	

This table presents 2010 summary statistics for hospitals above and below the median 2011 audit rate for two samples: all hospitals ("Overall Hospitals") and hospitals within 100 miles of the RAC border that have at least 1 hospital their "neighboring hospital comparison group" ("Border Hospitals"). Standard deviation is in parentheses. The median audit rate for the overall sample in 2011 was 1.78%. The median audit rate for border hospitals in 2011 was 1.60%. Bed size, urban status, and profit type status come from the Medicare Provider of Services file. Independent status comes from hospital merge 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).

Table 2. Coefficients of Across-Hospital Empirical Strategy

	(1)	(2)	(3)	(4)	(5)	(6)
	First Stage	Audit Demands	Medicare Admissions and Revenue		Hospital Compliance Costs	
	2011 Audit Rate	*Payment Demands ($1000s)*	*Log Admissions*	*Log Revenue*	*Log Admin Costs*	*Install Medical Necessity App.*
2011 jackknife state audit rate	1.0880*** (0.0666)					
2011 audit rate × 2011		314.1154*** (76.3316)	-0.0115** (0.0044)	-0.0102** (0.0053)	0.0154*** (0.0008)	0.0037 (0.0088)
2011 audit rate × 2012		179.5502 (115.1264)	-0.0192*** (0.0051)	-0.0177* (0.0080)	0.0068 (0.0012)	0.0217** (0.0079)
2011 audit rate × 2013		133.5120 (81.0827)	-0.0191** (0.0089)	-0.0280** (0.0092)	0.0034 (0.0012)	0.0225* (0.0129)
2011 audit rate × 2014		-5.1943 (54.6810)	-0.0113 (0.0114)	-0.0216 (0.0096)	0.0054 (0.0014)	0.0225* (0.0110)
2011 audit rate × 2015		53.9903** (22.6620)	-0.0193 (0.0148)	-0.0285 (0.0182)	-0.0014 (0.0107)	0.0090 (0.0123)
Hospital FE		X	X	X	X	X
Neighbor group FE		X	X	X	X	X
Hospitals	510 510	510	510	510	506	
N	510	46157	52139	52139	52107	36906
F		5.31	12.5	12.5	12.45	13.87

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the state and border segment level. This table reports the cross-sectional relationship captured in the first stage in Equation 1 and IV coefficients for 2011-2015 of the specification in Equation 2. Column 1 reports the first stage relationship between the 2011 jackknife state audit rate and audit rate. Columns 2-6 report the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome in a given year. Column 2 reports the effect on the amount of payment demanded from audits of prior admissions in thousands of dollars. Columns 3 and 4 report the effect on the log number of Medicare inpatient admissions and log Medicare inpatient revenue from the MEDPAR data. Column 5 reports the effect on log net administration costs from HCRIS data. Net administration costs are salary and other costs in the "administration and General" category in HCRIS, net of reclassifications and adjustments. Column 6 reports the effect on a dummy 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group." Omitted year is 2010.

Table 3. Summary statistics of patient characteristics by ED arrival time

	(1) Before MN		(2) After MN	
share inpatient	0.40		0.42	
share observation	0.05		0.05	
share white	0.78		0.77	
share Hispanic	0.12		0.11	
share female	0.57		0.54	
share inpatient in last 30 days	0.13		0.14	
share hospital visit in last 30 days	0.28		0.30	
average charges ($)	23966	(43649)	25881	(50655)
average age	68.04	(17.33)	68.22	(17.28)
average n of chronic conditions	3.95	(3.57)	4.17	(3.64)
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. "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 comes from HCUP SID and SEDD.

Table 4. Coefficients of After-Midnight ED Arrival Dummy on Patient Status and Outcomes

	(1)	(2)	(3)	(4)	(5)
	Medicare				Non-Medicare
	Inpatient	Observation	Out	Revisit 30d	Inpatient
$Post_{t(i)} \times AfterMN_{a(i)}$	-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 in parentheses and are clustered at the ED arrival hour and quarter level. This table reports coefficients on $Post_{t(i)} \times AfterMN_{a(i)}$ of the specification in Equation 3. $AfterMN_{a(i)}$ is a dummy if a visit occurred after midnight (midnight-3AM). $Post_{t(i)}$ is a dummy if a visit occurs after 2013Q1. "Admitted" is a dummy for whether the patient was eventually admitted as inpatient from the ED. "Observation" is a dummy for whether the patient was placed in observation status and was never admitted. "Out" is a dummy equal to one when a patient is neither admitted nor placed in observation status. "Revisit within 30 days" is a dummy 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. Sample for column 4 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 dummy, point of origin dummy, last ED visit within 30 days dummy, number of chronic conditions, and zip code income.

Table 5. Welfare Analysis Parameters

Case	(1) Baseline	(2) No compliance costs	(3) No effect on admissions
		Estimates	
Revenue effect θ_t	2011-2015: estimates after 2015: 2015 estimate	2011-2015: estimates after 2015: 2015 estimate	all years: 0
Comp. cost effect γ_t	2011-2015: estimates after 2015: 0	all years: 0	all years: 0
Demanded amt D_{a_t}	2011-2015: estimates after 2015: 0	all years: 0	all years: 0
2010 inpatient revenue	\$15,029,306	\$15,029,306	\$15,029,306
2010 compliance cost	\$12,822,887	\$12,822,887	\$12,822,887
		Parameters	
RAC contingency fee	10.75%	10.75%	10.75%
Marginal value of public funds	1.3	1.3	1.3
Discount rate	2%	2%	2%
Share of demanded pmts refunded	68%	68%	68%

This table lists the parameters and assumptions for the three welfare calculations depicted in Figure 7 and described in Section B. θ_t is the effect of a one percentage point increase in audit rate on hospital inpatient revenue, and the estimated coefficients before 2015 are from Table 2 column 3. γ_t is the effect of a one percentage point increase in audit rate on hospital compliance costs, and the estimated coefficients before 2015 are from Table 2 column 4. D_{a_t} is the effect of a one percentage point increase in audit rate on the amount demanded from audits and the estimated coefficients before 2015 are from Table 2. Cases 1 and 2 assume that θ_t is equal to θ_{2015} when $t > 2015$. The 2010 hospital revenue and hospital compliance costs (administration costs) are the median values for hospitals in the sample for Table 2.

A Additional Policy Context

A.1 Medicare Inpatient Prospective Payment System and Short Stays

CMS pays for inpatient hospital admissions through the inpatient prospective payment system (IPPS), in which CMS pays a fixed amount per inpatient stay within broad categories of diagnoses called Medicare Severity Diagnoses 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 incentive to keep costs low. The payment rate for each DRG reflects the national average cost of treating a patient across all cases, and 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 severity. It is also adjusted by hospital-specific factors like 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).

One persistent issue with IPPS, however, is the high number of short, unnecessary stays. CMS states 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, 2011). One less intensive alternative to an inpatient stay would be an outpatient observation stay, which consists of short-term (often diagnostic) services provider at the hospital while a physician decides whether to admit a patient or send them home. Observation stays typically last less than 48 hours and are billed as an outpatient service. The use of observation stays among Medicare beneficiaries has been growing rapidly over time (Medicare Payment Advisory Commission, 2015). An outpatient observation stay often precludes Medicare coverage for post-acute care services at a skilled nursing facility (SNF), because Medicare requires at least a three-day inpatient stay in order to cover a SNF stay.

From the patient's point of view, it is often difficult to differentiate between an observation stay and a short inpatient stay (Jaffe, 2016), so 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 admitting a patient for a short inpatient stay than for an observation stay – among DRGs common to both inpatient and observation stays, Medicare payments for inpatient stays were 2-3 times higher than payments for observation stays (Medicare Payment Advisory Commission, 2015).

CMS considered various alternative policy solutions to address unnecessary inpatient stays. CMS was wary of setting overly stringent admission requirements – CMS's admission guidelines give a lot of deference to physicians in the admission decision. CMS recognized the admission decision as a complex one, noting that providers must take into account many factors, including the "medical predictability of something adverse happening to the patient, the severity of the patient's condition, the need for and availability of diagnostics, the types of facilities available, hospital by-laws and admissions policies, and the relative appropriateness of treatment in each setting" (Centers for Medicare and Medicaid Services, 2012). CMS was also wary of reducing the payment rate for short stays or penalizing high rates of short stays, as policymakers were concerned hospitals would simply keep patients for longer to evade the policy (Medicare Payment Advisory Commission, 2015). There is evidence that hospitals delay discharging patients if there is incentive to do so (Jin et al., 2018). Finally, short stays (0-2 day stays) comprise almost a third of inpatient stays. the prevalence of short stays 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 Audit Process

RACs conducted post-payment reviews to identify and correct overpayments or underpayments for inpatient, outpatient, long term care, and durable medical equipment claims in the last three years. Figure A1 illustrates the claims auditing and appeals 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/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 CMS policies. 95 percent of audits were "complex reviews" in which a medical professional (e.g., 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's judgment to determine whether an overpayment or underpayment was made. Once the complex review is finished, RACs send a demand letter to providers which outlines whether a payment error was identified, the amount of overpayment or underpayment demanded, and references supporting the decision. 57 percent of complex reviews in 2011 result in no finding, 37 percent result in an overpayment demand (in which providers must return payment back to CMS), and six percent result in an underpayment demand (in which CMS returns payment to the provider). Hospitals can appeal demands by first requesting a redetermination by the RAC, and then escalating it to higher levels of appeals like requesting a reconsideration by a separate contractor or taking the appeal to court.

A.3 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 2005-2008 in select states, the RAC program was implemented nationally in 2010. At first, RACs were authorized only to audit claims with complex coding issues and for DRG validation. Each year, CMS expanded the scope of RAC audits, and in 2011 it expanded the scope to include to medical necessity reviews of inpatient claims. As shown in Figure 1, RAC audit activity peaked in 2011–2013, but 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 CMS over RAC audits[8]. 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 800,000 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, elevent percent of hospitals reported costs associated with managing the RAC program of over $100,000 (American Hospital Association, 2014)

Hospitals filed several complaints to CMS stating that RAC audits were overly aggressive. As a result, in 2014 CMS paused almost all RAC audits by significantly limiting their scope. Other Medicare contractors like MACs picked up auditing after the RACs were paused[9]. CMS maintained that the pause on RAC audits was temporary and would resume at previous levels, but it is clear from Figure 1 that RAC auditing never resumed to its peak levels after the pause in 2014. 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 provider burden. In August 2014, CMS also announced a one-time option to settle appeals by offering hospitals 68 percent of the 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

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

[9]For example, MACs conducted a program called "Teach, Probe, and Educate" in which they targeted 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 CMS for further steps.

350,000 appeals in exchange for \$1.5 billion in settled denials.

B Welfare Analysis Calculations

I next lay out the estimates required to calculate the marginal welfare effect in each year. Define θ_t and D_{at} as the estimates on log inpatient revenue and amount demanded in Table 2 and γ_t as the estimates on hospital administration costs in Table 2. Let I_{2010} be a hospital's inpatient revenue in 2010 (Table 5). Define I_{a_T} to be the present-discounted value of the marginal reduction in log inpatient revenue between 2011 and year T due to an exogenous increase in audit rate in 2011, relative to 2010. If θ_t is the estimated percent reduction in revenue in year t relative to 2010 (i.e., Table 2 column 4) and δ is the discount rate, then:[10]

$$I_{a_T} = \sum_{t=2011}^{T} \frac{\theta_t R_{2010}}{(1+\delta)^{t-2010}} \tag{10}$$

The total effect on government revenue also includes the money demanded back from audits D_{a_t} (i.e., Table 2 column 2) less the contingency fee f paid to RACs, which ranges from 9 to 12.5 percent of the amount demanded, and scaled by the share s of demanded payments that was refunded to hospitals in later lawsuits. For the main calculations 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 f is equal to the social administrative costs, otherwise f is an upper bound. Define $-R_{a_T}$ to be the present-discounted value of the marginal savings to the government of increasing the 2011 audit rate.[11]

$$-R_{a_T} = -I_{a_T} + (1-s) \sum_{t=2011}^{T} \frac{(1-f)D_{a_t}}{(1+\delta)^{t-2010}} \tag{11}$$

For provider compliance costs, let k_{2010} be a hospital's 2010 compliance costs (Table 5, k_{a_T} be the present-discounted value of the marginal increase in compliance costs between 2011 and year T, and γ_t be the estimated percent increase in compliance costs in year t relative to 2010 (i.e., Table 2 column 5). Then:

$$k_{a_T} = \sum_{t=2011}^{T} \frac{\gamma_t k_{2010}}{(1+\delta)^{t-2010}} \tag{12}$$

The marginal effect on government administrative costs m_{a_t} is defined as the contingency

[10] I_{a_T} is a negative number because θ_t is negative, and the marginal effect of increased auditing on hospital inpatient revenue is negative.

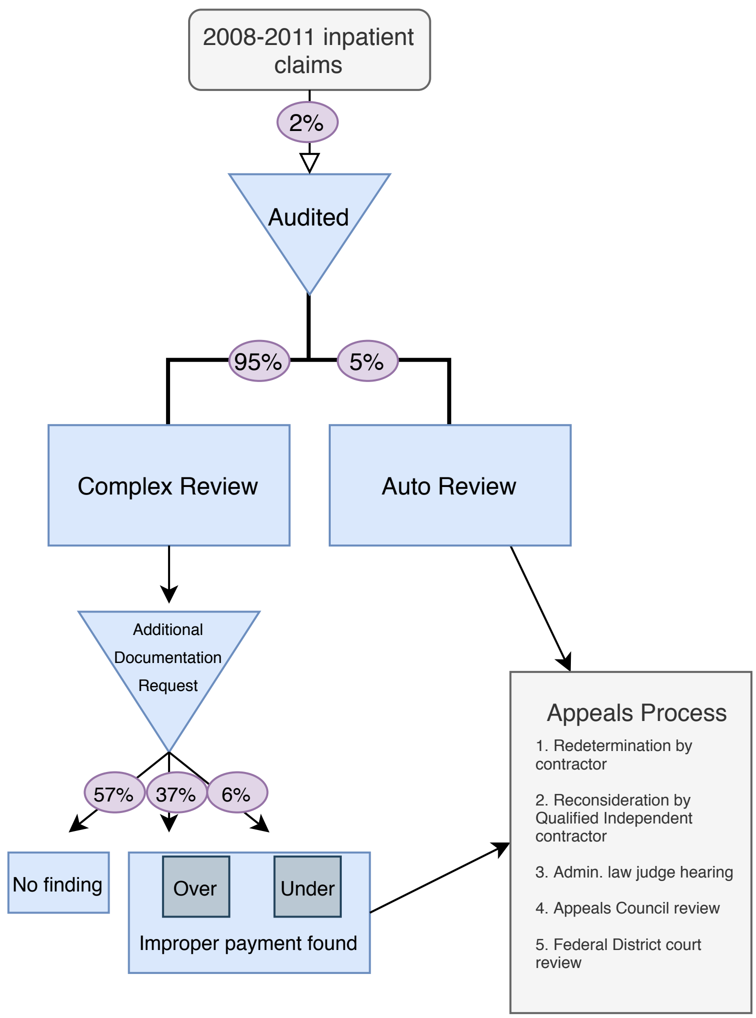
[11] $-R_{a_T}$ is a positive number.

fee f multiplied by the money demanded back from audits each year.

$$m_{a_T} = \sum_{t=2011}^{T} \frac{f D_{a_t}}{(1+\delta)^{t-2010}} \tag{13}$$

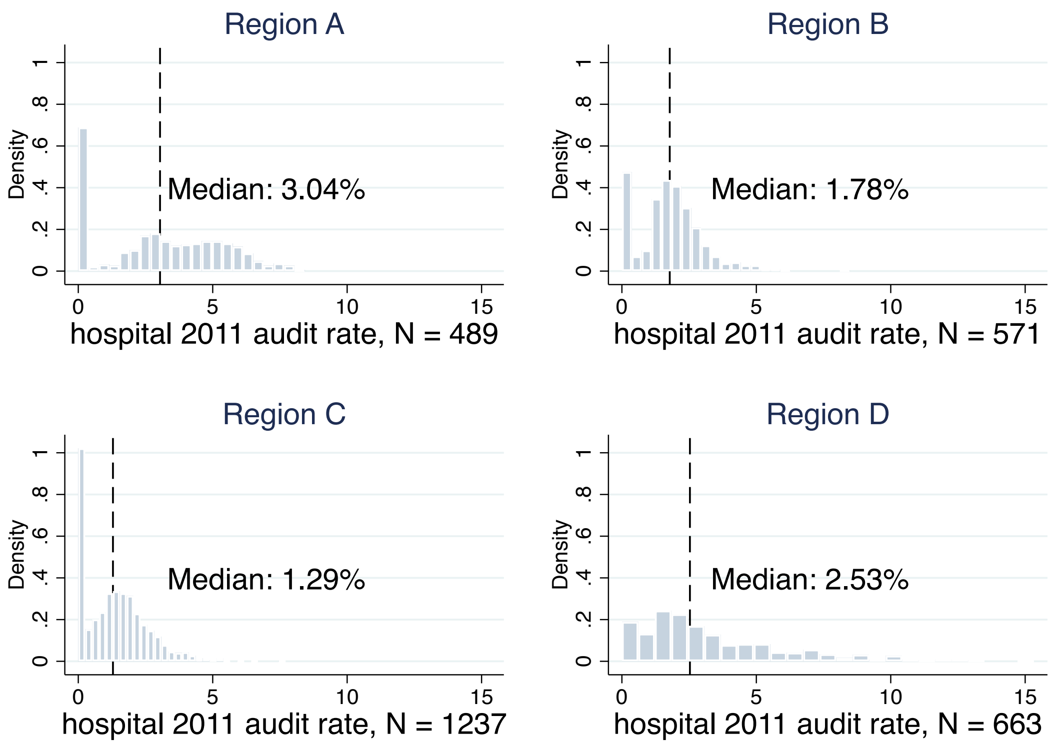
C Appendix Figures

Figure A1. RAC Inpatient Claims Auditing and Appeals Process



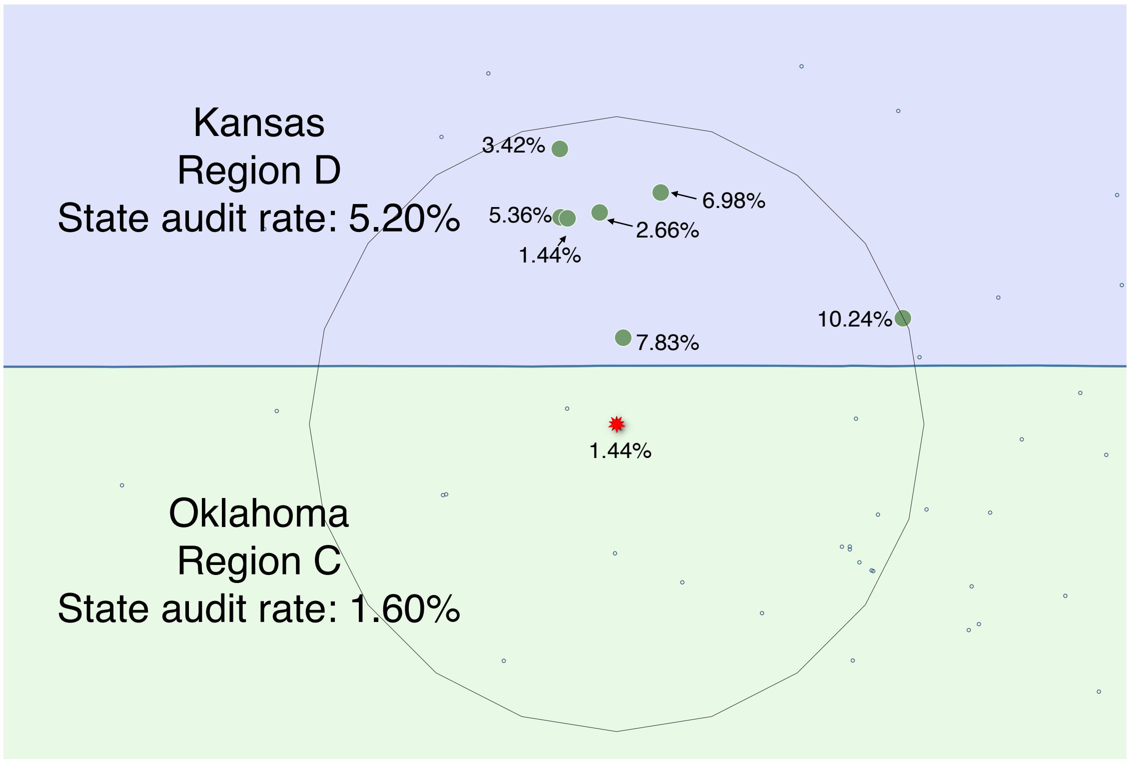
This figure illustrates the stages of the claims auditing and appeals process. The percentages in circles denote the percent of claims that, conditional on reaching a given stage in the process, reach the next stage. The percentages are calculated based on audits that occurred in 2011 of inpatient claims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

Figure A2. Histogram of 2011 Audit Rates by Region



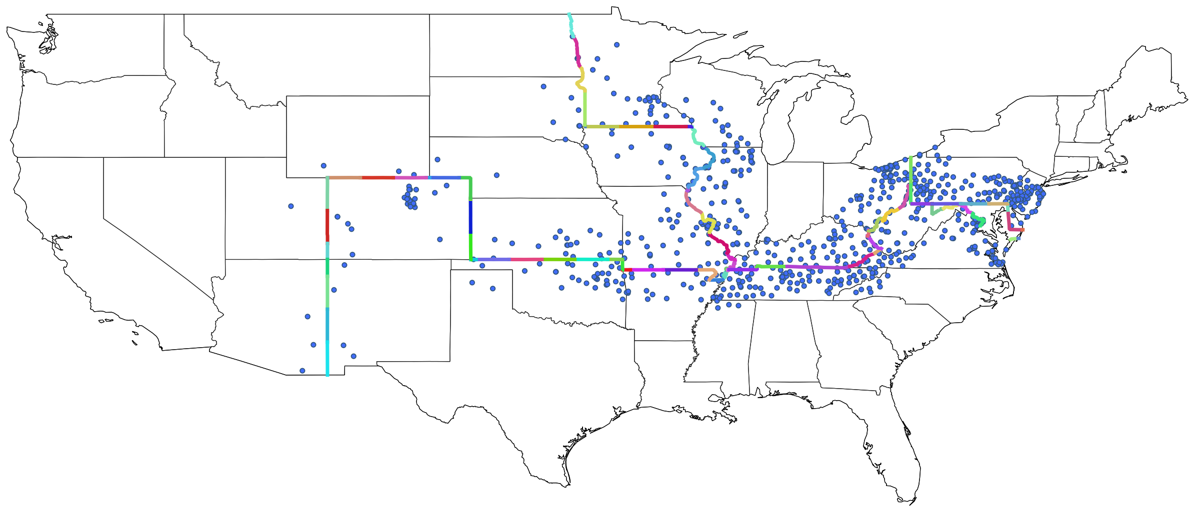
This figure plots the histogram of 2011 audit rates by RAC region, where audit rate is defined as the percent of a hospital's 2008-2011 claims that were audited by RACs.

Figure A3. Example of Border Hospital and Neighboring Hospital Sample Definition



This figure illustrates how "neighboring hospital 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 hospitals are considered neighboring hospitals to the red hospital.

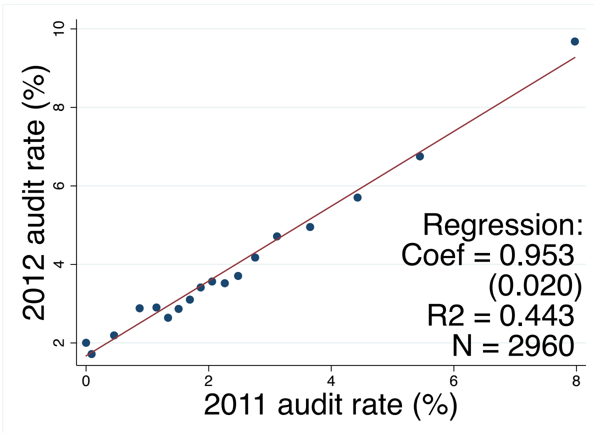
Figure A4. RAC Border Segments and Hospitals Within 100 Miles



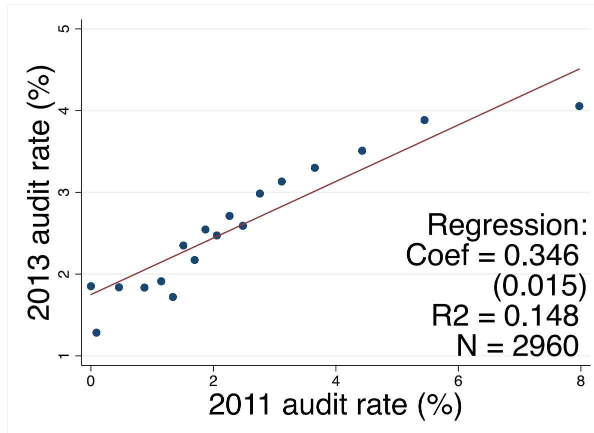
This figure shows how the RAC border is divided into 100 mile segments that do not cross state borders, and all hospitals within 100 miles of the RAC border.

Figure A5. Correlation of 2011 Audit Rate with Later Year Audit Rates

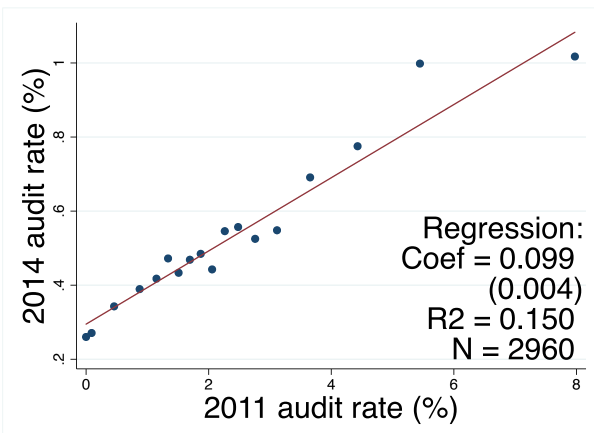
(a) 2011 vs. 2012 audit rates



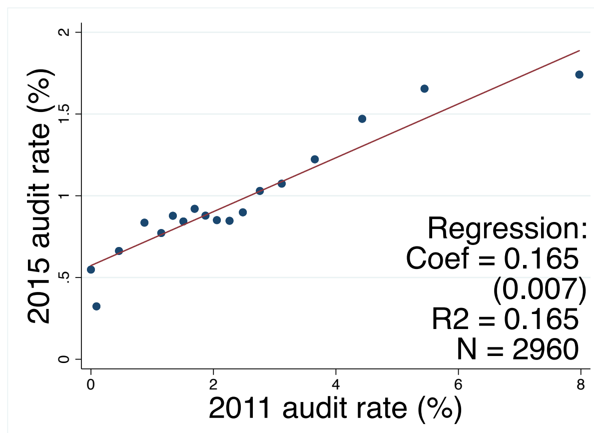
(b) 2011 vs. 2013 audit rates



(c) 2011 vs. 2014 audit rates



(d) 2011 vs. 2015 audit rates

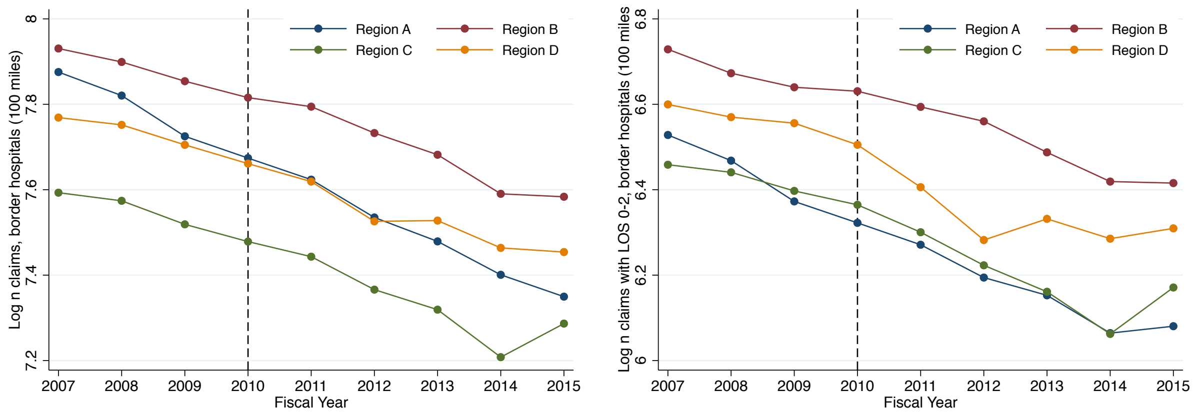


This figure plots binscatters of the correlation between hospital audit rates in 2011 and audit rates in subsequent years.

Figure A6. Medicare Admission Trends over Time by RAC region, Border Hospitals

(a) Log inpatient admissions

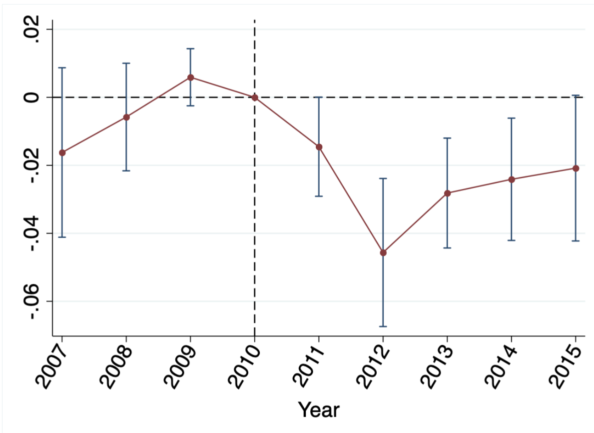
(b) Log inpatient admissions, LOS $\leq$ 2



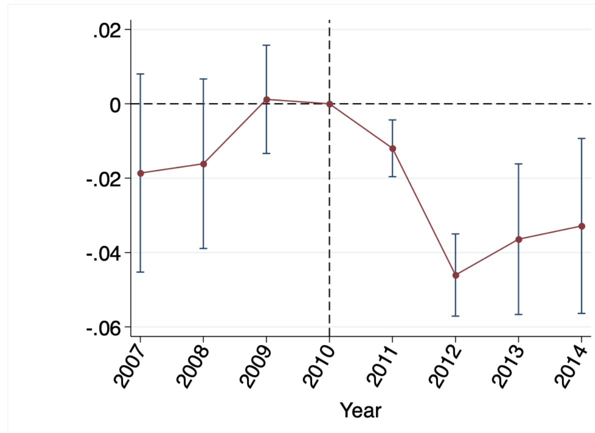
This figure plots Medicare admission trends over time among hospitals within 100 miles of the RAC border than have at least 1 hospital their "neighboring hospital comparison group."

Figure A7. Event Studies on Effect of 2011 Audit Rate on Medicare Admissions and Revenue

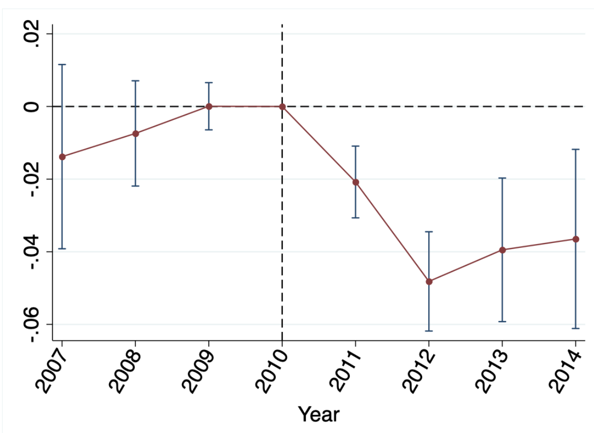
(a) Log Medicare admissions, LOS $\leq$ 2



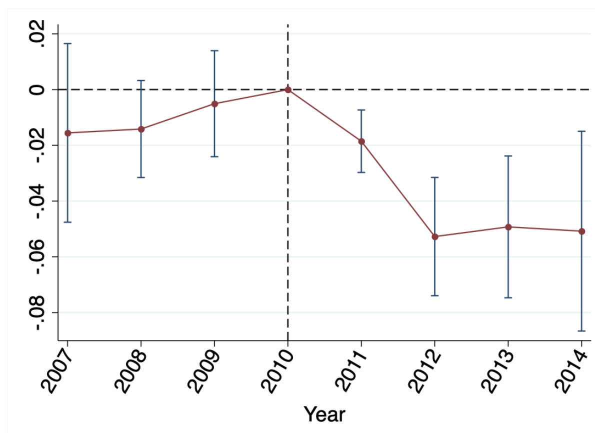
(b) Log Medicare inpatient revenue, LOS $\leq$ 2



(c) Log Medicare admissions, circulatory diagnoses



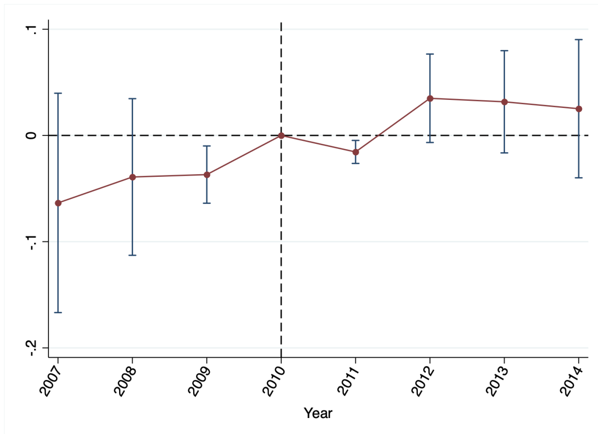
(d) Log Medicare inpatient revenue, circulatory diagnoses



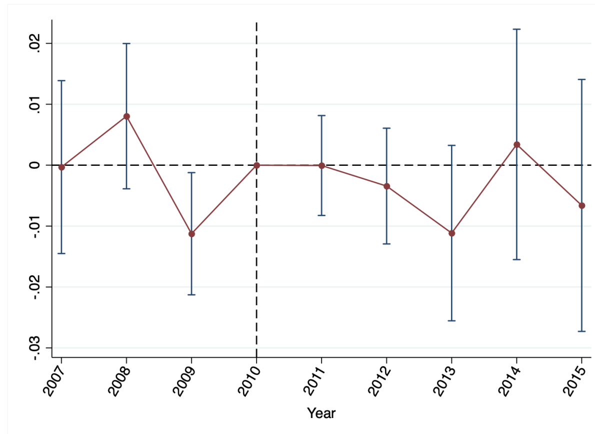
This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2. The omitted year is 2010. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome. Circulatory diagnoses are identified by identifying the Major Diagnostic Category associated with the DRG. 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group."

Figure A8. Event Studies of Effect of 2011 Audit Rate on Hospital Outpatient Revenue and Observation Stays

(a) Log observation stays

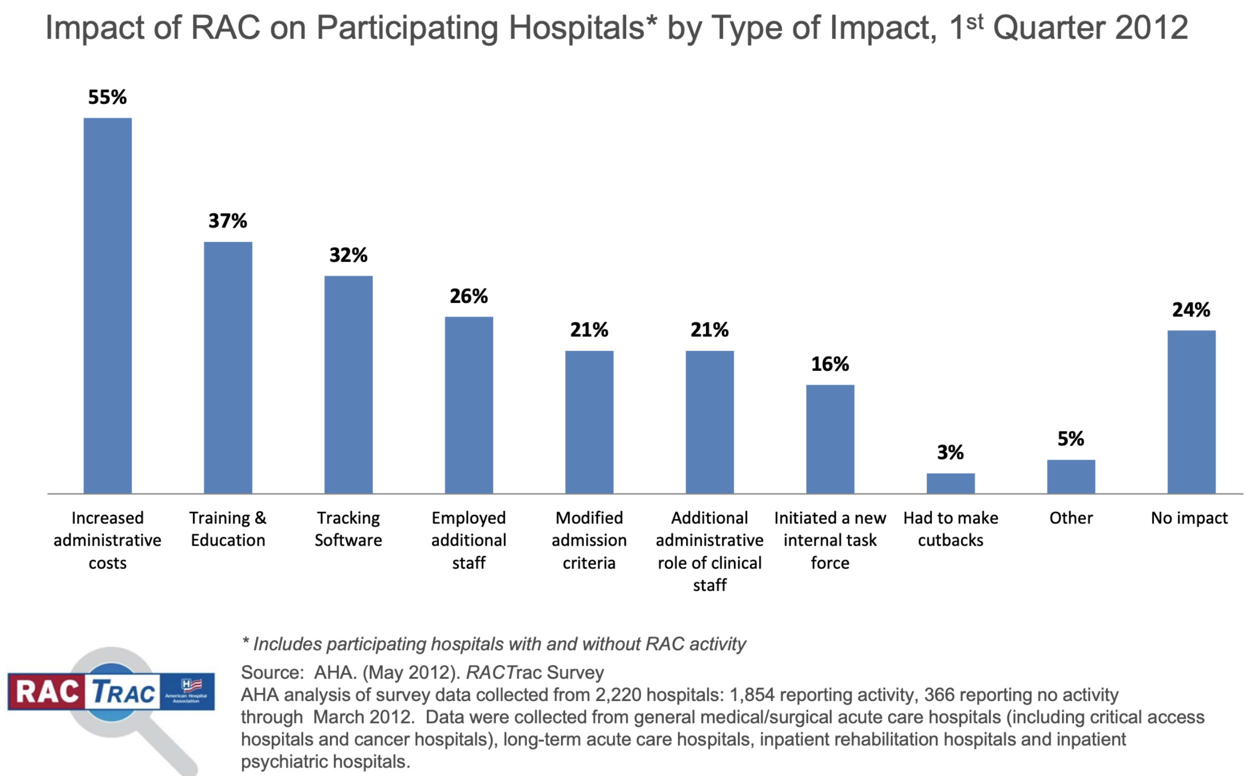


(b) Log Medicare outpatient revenue



This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2. The omitted year is 2010. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome in a given year. 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their “neighboring hospital comparison group.”

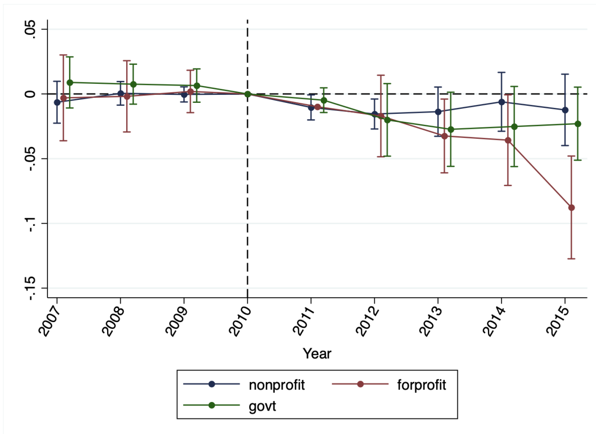
Figure A9. RACTrac Survey on Hospital Administration Spending, 2012 Quarter 1



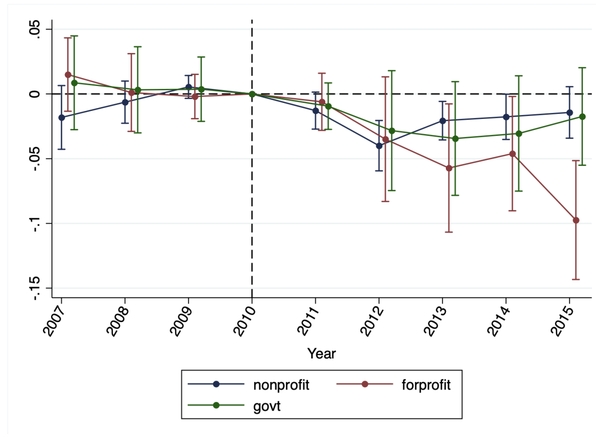
This figure is from a report published by the American Hospital Association on the RACTrac Survey, titled "Exploring the Impact of the RAC Program on Hospitals Nationwide" American Hospital Association (2012).

Figure A10. Event Studies on Heterogeneous Effects of 2011 Audit Rate on Medicare Admissions

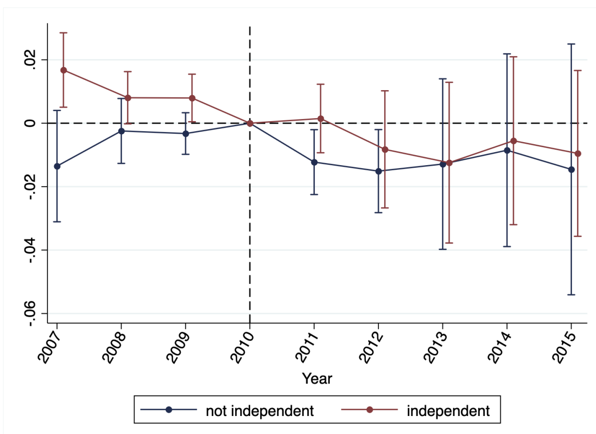
(a) Log adm., by hosp. profit type



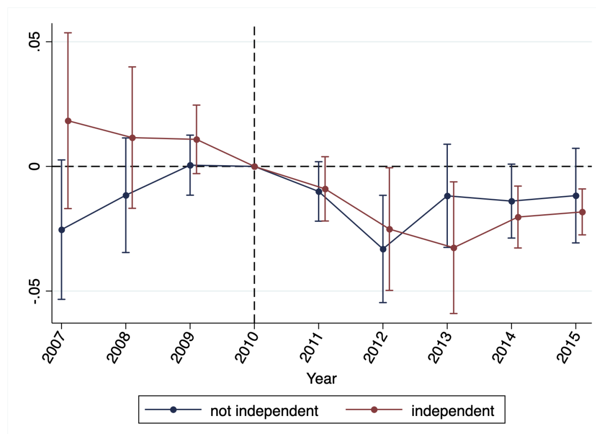
(b) Log adm., LOS $\leq$ 2, hosp. profit type



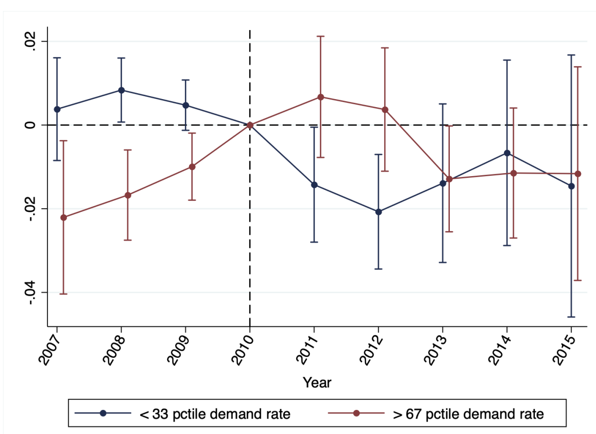
(c) Log adm., by independent status



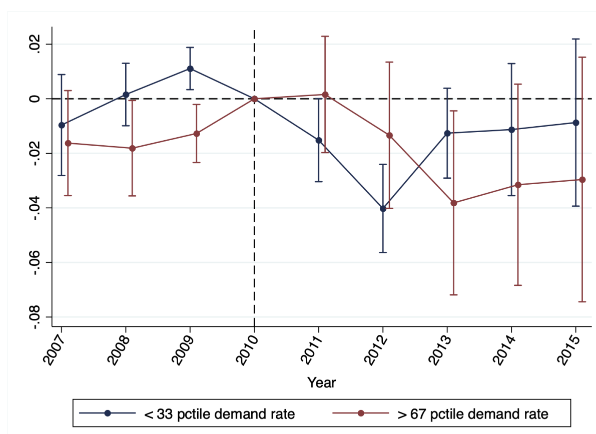
(d) Log adm., LOS $\leq$ 2, by ind. status



(e) Log adm., by demand rate



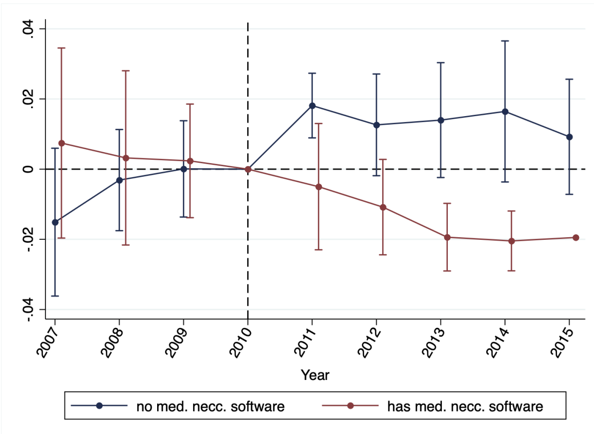
(f) Log adm., LOS $\leq$ 2, by demand rate



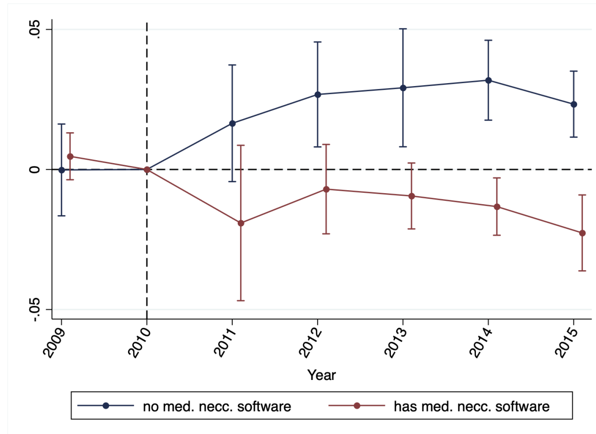
This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2, interacted with hospital type. The omitted year is 2010. "Hospital profit type" is derived from the Medicare Provider of Services file. "Independent" is defined in 2010 using the hospital merger data from (Cooper et al., 2019). "Demand rate" is the percent of audited claims that result in an overpayment/underpayment demand, and percentiles are relative to the distribution of demand rates in a hospital's state. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group."

Figure A11. Event Studies on Heterogeneous Effects of 2011 Audit Rate on Hospital Administrative Burden

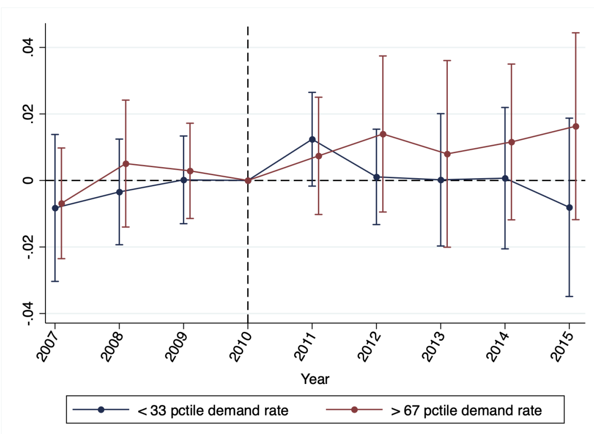
(a) Log admin costs, by 2010 med. necc. software status



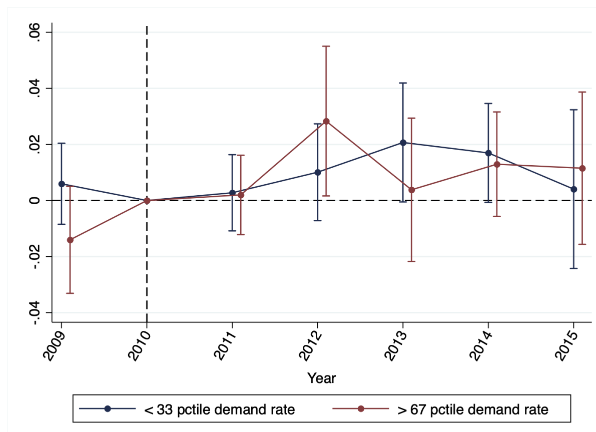
(b) Dummy for installing med. necc. software, by 2010 med. necc. software status



(c) Log admin costs, by demand rate



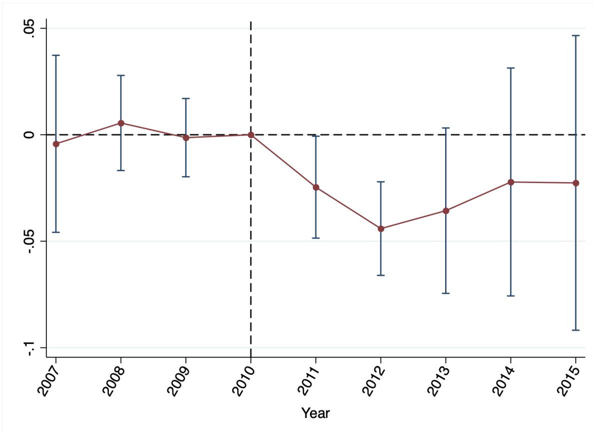
(d) Dummy for installing med. necc. software, by demand rate



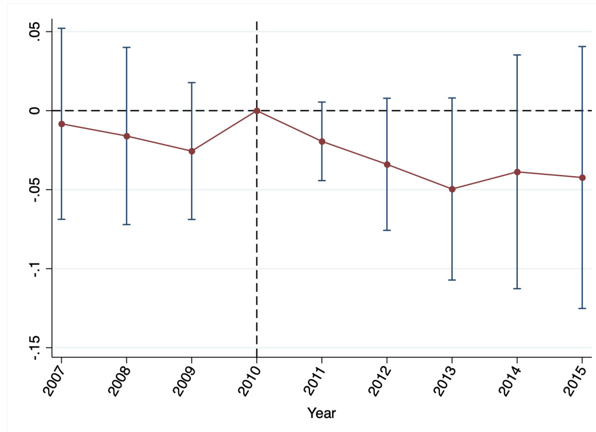
This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2, interacted with hospital type. The results are clustered at the state and border segment level. The omitted year is 2010. "2010 medical necessity software status" is based on whether a hospital has installed or was already contracted to install medical necessity software in 2010 in HIMSS. "Demand rate" is the percent of audited claims that result in an overpayment/underpayment demand, and percentiles are relative to the distribution of demand rates in a hospital's state. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group."

Figure A12. Event Studies on Effect of 2011 Denial Rate on Medicare Admissions and Revenue

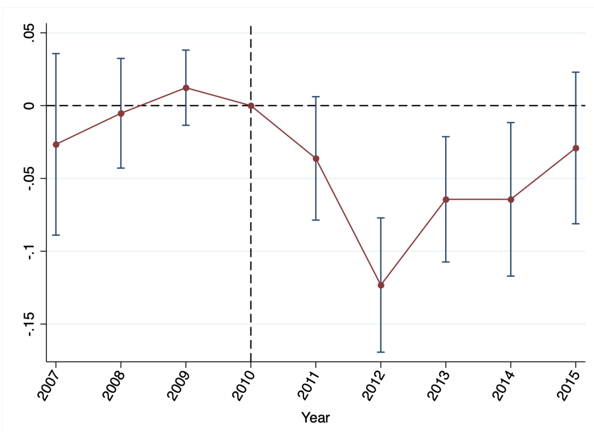
(a) Log Medicare admissions



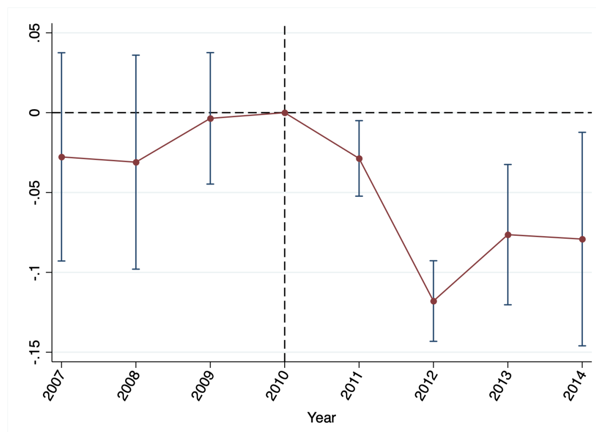
(b) Log Medicare inpatient revenue



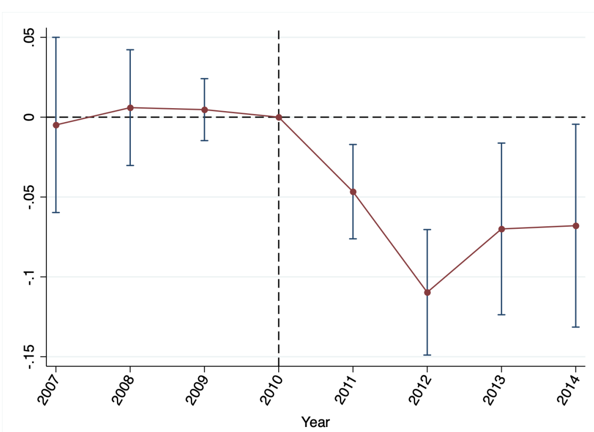
(c) Log Medicare admissions, LOS ≤ 2



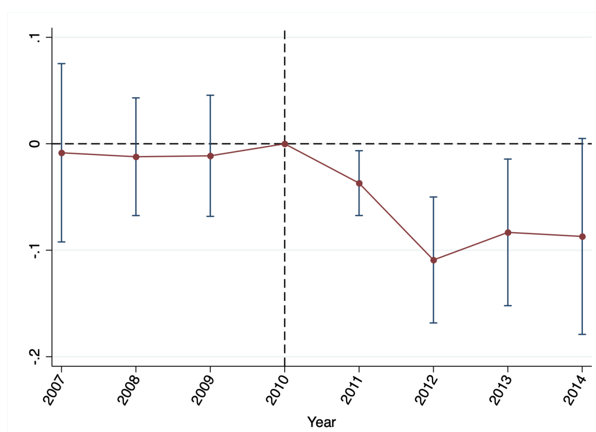
(d) Log Medicare inpatient revenue, LOS ≤ 2



(e) Log Medicare admissions, circulatory diagnoses



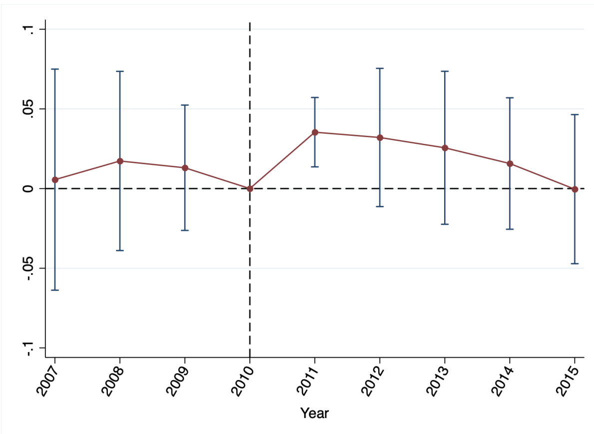
(f) Log Medicare inpatient revenue, circulatory diagnoses



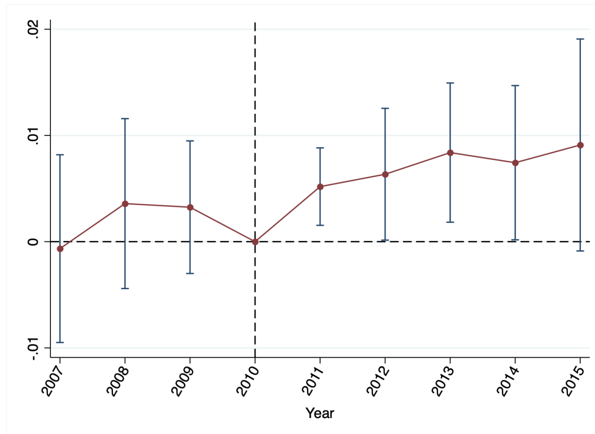
This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2. The results are clustered at the state and border segment level. The omitted year is 2010. Each coefficient estimates the effect of a 1pp increase in 2011 denial rate (share of claims that are audited and result in an overpayment/underpayment demand) on a hospital-level outcome. Circulatory diagnoses are identified by identifying the Major Diagnostic Category associated with the DRG. 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their

Figure A13. Event Studies of Effect of 2011 Denial Rate on Hospital Administrative Burden

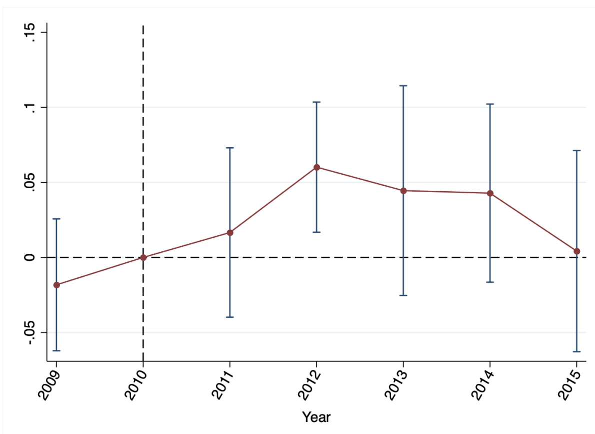
(a) Log net administration costs



(b) Administration cost share

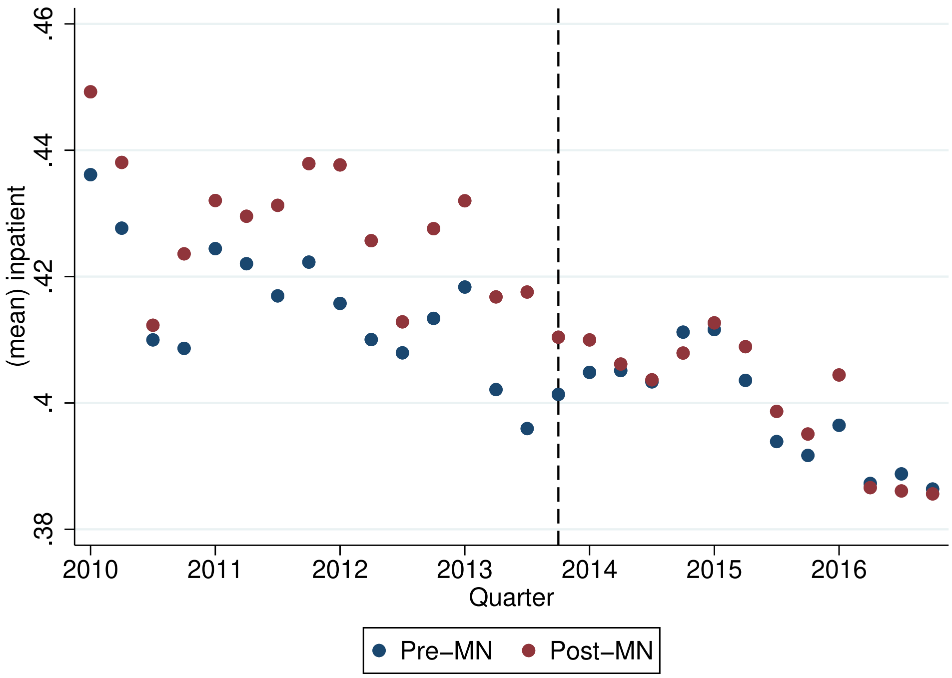


(c) Dummy for installing medical necessity software



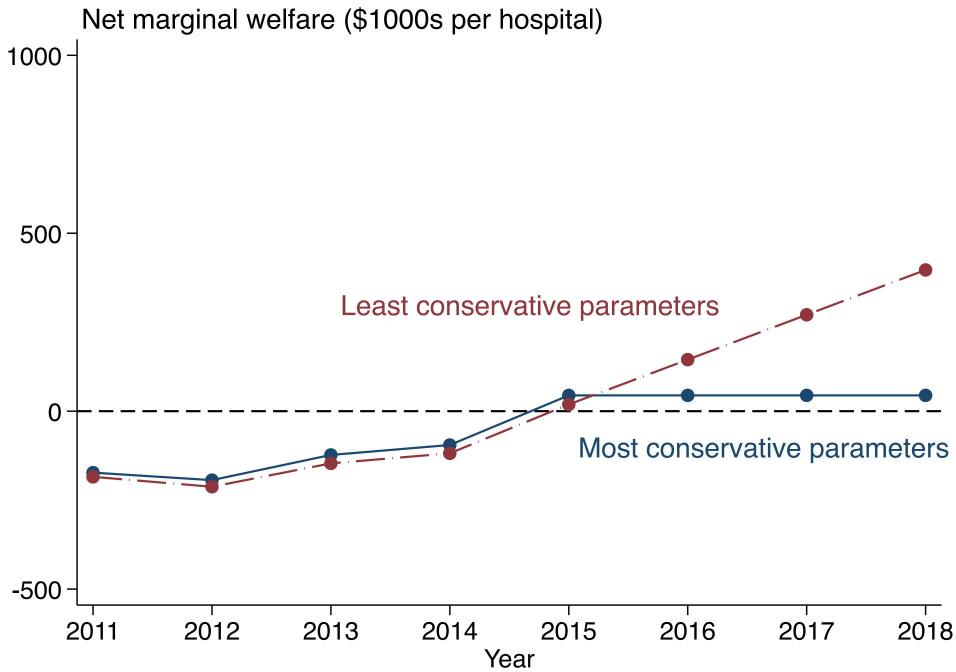
This figure plots event studies of the IV coefficients and 95% confidence intervals of the specification in Equation 2. The results are clustered at the state and border segment level. The omitted year is 2010. Each coefficient estimates the effect of a 1pp increase in 2011 denial rate (share of claims that are audited and result in an overpayment/underpayment demand) on a hospital-level outcome in a given year. Net administration costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Administration cost share is defined as net administration costs divided by net total costs for a hospital. Dummy 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group."

Figure A14. Share of Medicare ED Patients Admitted as Inpatient, Before- vs. After-Midnight ED Arrivals



This figure plots the average share of patients admitted among before-midnight and after-midnight ED arrivals by quarter, among traditional Medicare patients who arrived in the ED within 3 hours of midnight in Florida.

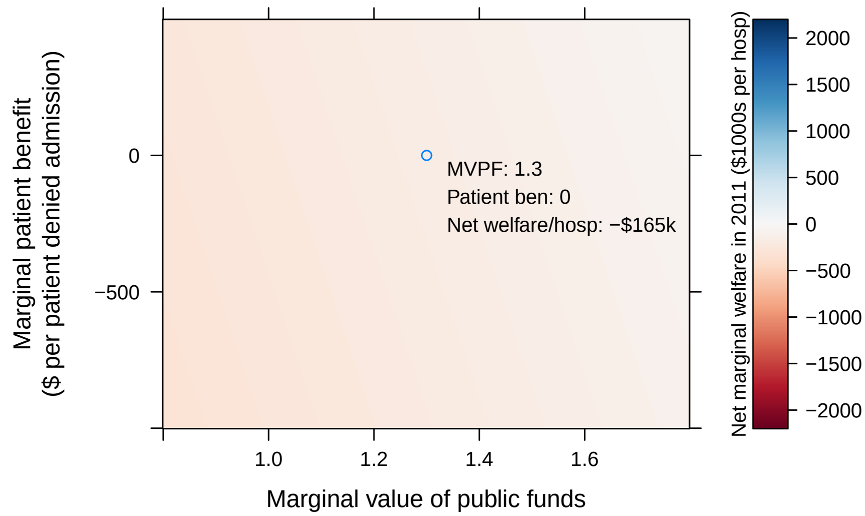
Figure A15. Welfare Analysis Estimates, Most vs. Least conservative



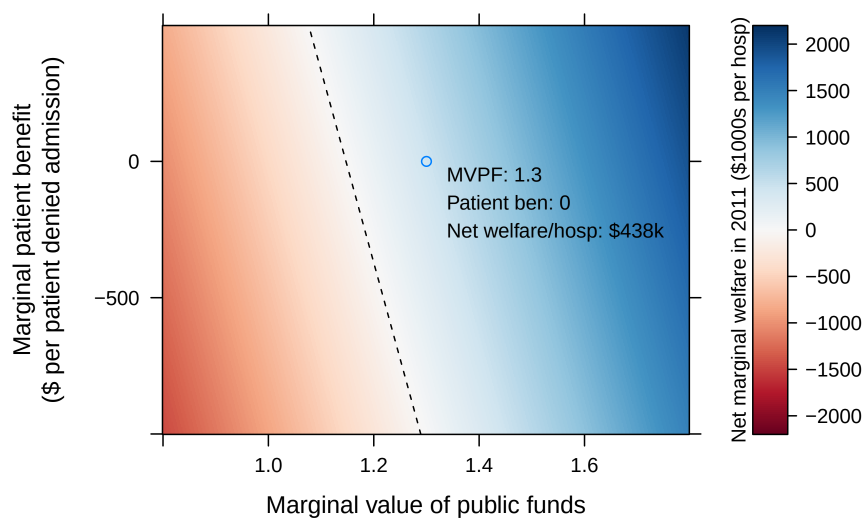
This figure plots the per-hospital welfare effect, or the difference between the savings and compliance/administrative costs of auditing, of increasing audits in 2011 by a given year. Increasing audits is welfare-improving if this value is positive and welfare-reducing if this value is negative. The figure plots the estimates from the most conservative case to the least conservative case. In the most conservative case, the RACs charge the highest contingency fee of 12.5 percent, the effect on admissions is 0 after 2015, and CMS has to refund 68 percent of demanded payments. In the least conservative case, RACs demand the lowest contingency fee of 9 percent, the effect on admissions is permanent after 2015, and CMS keeps all the demanded payments.

Figure A16. Marginal Welfare Effect by Patient Benefit Marginal Value of Public Funds, 2011 and 2018

(a) 2011



(b) 2018



This figure plots the per-hospital marginal welfare effect of increasing 2011 audits, with varying assumptions about the marginal value of public funds and the marginal patient benefit ($ per patient denied admission) in (a) 2011 and (b) 2018. Increasing audits is welfare-improving if this value is positive (blue) and welfare-reducing if this value is negative (red). Table 5 lists the other parameters and estimates used to calculate the welfare effects. The blue point represents the baseline specification, which assumes a MVPF of 1.3 and no patient health effects from reduced admissions. The dotted line denotes the set of combinations of marginal patient benefit and marginal value of public funds where the marginal welfare effect is 0.

D Appendix Tables

Table B1. Summary Statistics of 2010 Hospital Characteristics by 2011 Hospital Audit Rate

	(1) Above median 2011 audit rate		(2) Below median 2011 audit rate	
2011 audit rate (%)	3.60	(1.89)	0.73	(0.65)
beds	182.04	(164.09)	228.65	(195.55)
share urban	0.68		0.76	
share in region A	0.23		0.11	
share in region B	0.19		0.20	
share in region C	0.28		0.54	
share in region D	0.30		0.16	
total costs (million \$)	193.78	(248.46)	214.77	(269.23)
net admin costs (million \$)	28.84	(39.11)	32.08	(39.62)
mean admin share	0.15	(0.05)	0.16	(0.04)
wage index	0.99	(0.17)	0.97	(0.16)
mean m'care reimb/enrollee (\$)	9392	(1370)	9960	(1438)
mean hospital/snf reimb/enrollee (\$)	4579	(764)	4769	(727)
m'care admissions	3057	(3058)	3927	(3351)
mean m'care length of stay	4.64	(0.96)	4.88	(0.96)
mean m'care inpatient payment (\$)	8789	(3135)	8998	(3100)
total m'care inpatient rev. (million \$)	30.28	(38.07)	38.95	(42.14)
mean share claims LOS 0-2	0.31	0.07	0.30	(0.07)
mean share claims circulatory diagnosis	0.19	(0.06)	0.21	(0.08)
Observations	1474		1427	

This table presents 2010 summary statistics for hospitals above and below the median 2011 audit rate. The overall median audit rate in 2011 was 1.92%. Bed size and urban status come from the Medicare Provider of Services file. Total and administrative costs come from HCRIS. Wage index comes from the Medicare impact file. Average Medicare reimbursement and hospital/SNF reimbursement per enrollee are based on county measures in the Dartmouth Atlas and are adjusted for demographics and prices. Medicare volume and inpatient stay characteristics are from MEDPAR. Mean inpatient characteristics are defined as the average of each hospital's average (i.e., mean medicare LOS is weighted by number of hospitals, not number of claims).

Table B2. Across-Hospital Reduced Form and IV Coefficients, Medicare Admissions and Revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	2SLS						Reduced Form					
	Overall		LOS ≤ 2		Circulatory		Overall		LOS ≤ 2		Circulatory	
	Log Adm.	*Log. Rev.*	*Log Adm.*	*Log. Rev.*	*Log Adm.*	*Log. Rev.*	*Log Adm.*	*Log. Rev.*	*Log Adm.*	*Log. Rev.*	*Log Adm.*	*Log. Rev.*
2011 audit rate	-0.0115**	-0.0102**	-0.0145*	-0.0120***	-0.0208***	-0.0185***						
× 2011	(0.0044)	(0.0044)	(0.0074)	(0.0039)	(0.0050)	(0.0057)						
2011 audit rate	-0.0192***	-0.0177*	-0.0457***	-0.0461***	-0.0482***	-0.0527***						
× 2012	(0.0051)	(0.0093)	(0.0111)	(0.0056)	(0.0070)	(0.0108)						
2011 audit rate	-0.0191**	-0.0280**	-0.0282***	-0.0364***	-0.0395***	-0.0493***						
× 2013	(0.0089)	(0.0129)	(0.0082)	(0.0103)	(0.0101)	(0.0130)						
2011 audit rate	-0.0113	-0.0216	-0.0241**	-0.0329**	-0.0365***	-0.0508**						
× 2014	(0.0114)	(0.0157)	(0.0092)	(0.0120)	(0.0126)	(0.0183)						
2011 audit rate	-0.0193	-0.0285	-0.0208*									
× 2015	(0.0148)	(0.0182)	(0.0109)									
2011 jackknife state audit rate							-0.0119**	-0.0105*	-0.0150**	-0.0123***	-0.0214***	-0.0191**
× 2011							(0.0046)	(0.0053)	(0.0071)	(0.0041)	(0.0052)	(0.0072)
2011 jackknife state audit rate							-0.0198***	-0.0182*	-0.0470***	-0.0473***	-0.0496***	-0.0543***
× 2012							(0.0054)	(0.0101)	(0.0084)	(0.0036)	(0.0076)	(0.0127)
2011 jackknife state audit rate							-0.0197*	-0.0287*	-0.0289***	-0.0374***	-0.0406***	-0.0506***
× 2013							(0.0096)	(0.0143)	(0.0062)	(0.0093)	(0.0111)	(0.0153)
2011 jackknife state audit rate							-0.0116	-0.0222	-0.0248***	-0.0337***	-0.0375**	-0.0522**
× 2014							(0.0120)	(0.0168)	(0.0088)	(0.0119)	(0.0143)	(0.0205)
2011 jackknife state audit rate							-0.0198	-0.0292	-0.0214*			
× 2015							(0.0162)	(0.0201)	(0.0112)			
Hosp FE	X	X	X	X	X	X	X	X	X	X	X	X
Nbr group FE	X	X	X	X	X	X	X	X	X	X	X	X
Hosp	510	510	510	510	510	510	510	510	510	510	510	510
N	52139	52139	52139	46437	46453	46453	52139	52139	52139	46437	46453	46453
F	12.5	12.5	12.5	15.27	15.27	15.27						

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the state and border segment level. This table reports reduced form IV coefficients for 2011-2015 of the specification in Equation 2. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome in a given year. Columns 1-6 report the two stage least squares outcomes for the number of and revenue from Medicare admissions overall (columns 1-2), with length of stay ≤ 2 (columns 3-4), and the number of and revenue from circulatory admissions (columns 5-6). Columns 7-12 report the reduced form coefficients from $Jackknife\ Audit\ Rate_{hs(h),2011} \times Year_y$ from Medicare admissions overall (columns 7-8), with length of stay ≤ 2 (columns 9-10), and the number of and revenue from circulatory admissions (columns 11-12). Circulatory diagnoses are identified by the Major Diagnostic Category associated with the DRG. 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group." Omitted year is 2010.

Table B3. Across-Hospital Reduced Form and IV Coefficients, Medicare Admissions and Revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2SLS				Reduced Form			
	LOS > 2		Non-circulatory		LOS > 2		Non-circulatory	
	Log Adm.	*Log. Rev.*	*Log Adm.*	*Log. Rev.*	*Log Adm.*	*Log. Rev.*	*Log Adm.*	*Log. Rev.*
2011 audit rate × 2011	-0.0092	-0.0102**	-0.0084*	-0.0068*				
	(0.0064)	(0.0044)	(0.0041)	(0.0038)				
2011 audit rate × 2012	-0.0077	-0.0178*	-0.0121**	-0.0080				
	(0.0091)	(0.0093)	(0.0044)	(0.0080)				
2011 audit rate × 2013	-0.0145	-0.0280**	-0.0136	-0.0221*				
	(0.0118)	(0.0129)	(0.0083)	(0.0127)				
2011 audit rate × 2014	-0.0049	-0.0223	-0.0050	-0.0138				
	(0.0133)	(0.0154)	(0.0107)	(0.0148)				
2011 audit rate × 2015	-0.0166		-0.0154					
	(0.0176)		(0.0145)					
2011 jackknife state audit rate × 2011					-0.0095	-0.0105*	-0.0087*	-0.0070
					(0.0068)	(0.0053)	(0.0044)	(0.0045)
2011 jackknife state audit rate × 2012					-0.0079	-0.0183*	-0.0125**	-0.0083
					(0.0097)	(0.0100)	(0.0045)	(0.0086)
2011 jackknife state audit rate × 2013					-0.0149	-0.0287*	-0.0140	-0.0227
					(0.0128)	(0.0143)	(0.0089)	(0.0139)
2011 jackknife state audit rate × 2014					-0.0050	-0.0229	-0.0051	-0.0142
					(0.0139)	(0.0165)	(0.0111)	(0.0157)
2011 jackknife state audit rate × 2015					-0.0170		-0.0158	
					(0.0190)		(0.0157)	
Hosp FE	X	X	X	X	X	X	X	X
Nbr group FE	X	X	X	X	X	X	X	X
Hosp 510	510	510	510	510	510	510	510	510
N 510	52139	46437	52139	46453	52139	46437	52139	46453
F		12.5	15.27	12.5	15.27			

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the state and border segment level. This table reports reduced form IV coefficients for 2011-2015 of the specification in Equation 2. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome in a given year. Columns 1-4 report the two stage least squares outcomes for the number of and revenue from Medicare admissions with length of stay > 2 (columns 1-2) and the number of and revenue from non-circulatory admissions (3-4). Columns 5-8 report the reduced form coefficients from $Jackknife\ Audit\ Rate_{hs(h),2011} \times Year_y$ from Medicare admissions with length of stay > 2 (columns 5-6) and the number of and revenue from non-circulatory admissions (columns 7-8). Circulatory diagnoses are identified by the Major Diagnostic Category associated with the DRG. 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group." Omitted year is 2010.

Table B4. Across-Hospital Reduced Form and IV Coefficients, Hospital Administrative Burden

	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS			Reduced Form		
	Administration Costs		Software Installation	Administration Costs		Software Installation
	Log Costs	*Share*	*Medical Necc.*	*Log Costs*	*Share*	*Medical Necc.*
2011 audit rate × 2011 (0.0067)	0.0191*** (0.0009)	0.0021** (0.0149)	0.0008			
2011 audit rate × 2012	0.0065 (0.0115)	0.0017 (0.0019)	0.0246** (0.0113)			
2011 audit rate × 2013 (0.0120)	0.0010 (0.0019)	0.0025 (0.0195)	0.0204			
2011 audit rate × 2014	0.0031 (0.0128)	0.0024 (0.0022)	0.0210 (0.0166)			
2011 audit rate × 2015	-0.0038 (0.0121)	0.0039 (0.0026)	0.0061 (0.0166)			
2011 jackknife state audit rate × 2011				0.0205** (0.0078)	0.0023** (0.0010)	0.0006 (0.0153)
2011 jackknife state audit rate × 2012				0.0069 (0.0124)	0.0018 (0.0020)	0.0255** (0.0109)
2011 jackknife state audit rate × 2013				0.0010 (0.0131)	0.0027 (0.0021)	0.0210 (0.0204)
2011 jackknife state audit rate × 2014				0.0033 (0.0136)	0.0026 (0.0024)	0.0217 (0.0174)
2011 jackknife state audit rate × 2015				-0.0041 (0.0131)	0.0042 (0.0029)	0.0061 (0.0168)
Hosp FE	X	X	X	X	X	X
Nbr group FE	X	X	X	X	X	X
Hosp	498	498	494	498	498	494
N	51207	51239	36228	51207	51239	36228
F	14.55	14.55	16.29			

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered at the state and border segment level. This table reports reduced form IV coefficients for 2011-2015 of the specification in Equation 2. Each coefficient estimates the effect of a 1pp increase in 2011 audit rate on a hospital-level outcome in a given year. Columns 1-3 report the two stage least squares outcomes for log net administration costs (1), share administration costs (2), and a dummy for installation of medical necessity software (3). Columns 4-6 report the two stage least squares outcomes for log net administration costs, share administration costs, and a dummy for installation of medical necessity software. Net administration costs are salary and other costs in the "Administrative and General" category in HCRIS, net of reclassifications and adjustments. Administration cost share is defined as net administration costs divided by net total costs for a hospital. Dummy 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. Sample is comprised of hospitals within 100 miles of the RAC border with at least 1 hospital in their "neighboring hospital comparison group." Omitted year is 2010.

Table B5. After-Midnight ED Arrival, Within 1 Hour of Midnight vs. 4 Hours from Midnight

	(1) After Midnight$_{a(i)}$
$Post_{t(i)} \times$ (Within 1 Hr of MN)$_{a(i)}$	0.004
	(0.003)
Observations	1060962

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered by the ED arrival hour and quarter. This table reports estimates and standard errors of the $Post_{t(i)} \times$ Within 1 Hr of MN$_{h(i)}$ coefficient in Equation 6. Within 1 Hr of MN$_{h(i)}$ is a dummy equal to one for patients who arrive within an hour of midnight (11PM-1AM) and zero for patients who arrive 4 hours away from midnight (either 8PM-9PM or 3AM-4AM). After Midnight$_a(i)$ is a dummy for whether the ED arrival hour was after midnight. Regression includes hospital and hospital-quarter fixed effects.

Table B6. Within-hospital Difference-in Difference Results by Predicted Risk Decile

	(1) Inpatient	(2) Revisit 30d
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 1})_i$	0.015*** (0.003)	0.001 (0.003)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 2})_i$	-0.006** (0.002)	-0.002 (0.005)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 3})_i$	-0.018*** (0.004)	0.001 (0.005)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 4})_i$	-0.018*** (0.007)	0.009 (0.006)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 5})_i$	-0.052*** (0.008)	0.004 (0.006)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 6})_i$	-0.055*** (0.006)	-0.005 (0.007)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 7})_i$	-0.036** (0.011)	0.003 (0.007
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 8})_i$	-0.009 (0.014)	-0.008 (0.005)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 9})_i$	-0.007 (0.010)	-0.000 (0.004)
$\text{Post}_{t(i)} \times \text{AfterMN}_{a(i)}$ $\times(\text{Risk Decile 10})_i$	-0.003 (0.004)	-0.002 (0.005)
Observations	1236048	1236048

* $p < .10$, ** $p < .05$, *** $p < .01$. Standard errors in parentheses and are clustered by the ED arrival hour and quarter. This table reports estimates and standard errors of the $Post_{t(i)} \times AfterMN_{a(i)}$ coefficient in Equation 3, interacted with a dummy for predicted risk decile. Patient risk is predicted by estimating a logit using ED visits between 9AM and 3PM of a dummy 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 dummy, point of origin dummy, 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.

Table B7. Welfare Analysis Parameters, Robustness

Case	Most conservative	Least conservative
	Estimates	
Revenue effect θ_t	2011-2015: est. coeffs after 2015: 0	2011-2015: est. coeffs after 2015: 2015 estimate
Comp. cost effect γ_t	2011-2015: est. coeffs after 2015: 0	2011-2015: est. coeffs after 2015: 0
Demanded amt D_{a_t}	2011-2015: est. coeffs after 2015: 0	2011-2015: est. coeffs after 2015: 0
2010 hospital revenue	\$15,029,306	\$15,029,306
2010 hospital compliance costs	\$12,822,887	\$12,822,887
	Parameters	
RAC contingency fee	12.5%	9%
Value of public funds	1.3	1.3
Discount rate	2%	2%
Share of demanded pmts refunded	68%	0%

This table lists the parameters and assumptions for "most conservative" and "least conservative" calculations, depicted in Figure A15. θ_t is the effect of a one percentage point increase in audit rate on hospital inpatient revenue, and the estimated coefficients before 2015 are from Table 2 column 3. γ_t is the effect of a one percentage point increase in audit rate on hospital compliance costs, and the estimated coefficients before 2015 are from Table 2 column 4. D_{a_t} is the effect of a one percentage point increase in audit rate on the amount demanded from audits and the estimated coefficients before 2015 are from Table 2. The 2010 hospital revenue and hospital compliance costs (administration costs) are the median values for hospitals in the sample for Table 2.