Can Whistleblowers Root Out Public Expenditure Fraud?
Evidence from Medicare

Jetson Leder-Luis*

Abstract

This paper analyzes private anti-fraud enforcement under the False Claims Act, which compensates whistleblowers for litigating against healthcare providers who overbill the government. I conduct several case studies of successful whistleblower lawsuits concerning Medicare fraud, pairing new legal data with large samples of Medicare claims. I estimate that deterrence from $1.9 billion in whistleblower settlements generated Medicare cost savings of nearly $19 billion, while imposing low costs on the government. In a case study of fraudulent spine surgery, whistleblower-induced changes to care modestly improved patient health. These results suggest private enforcement is a cost-effective way to combat public expenditure fraud.
(JEL Codes: H41, K42, I13, D73)

*Boston University, Questrom School of Business, Department of Markets, Public Policy, and Law.
Mailing: 595 Commonwealth Avenue Office 518E, Boston, MA 02115
Email: jetson@bu.edu

1   Introduction

Waste, fraud and abuse are serious concerns in the governmental provision of goods and services. Governments often rely on private firms to execute their spending, such as the procurement of goods or in contracting to provide services, and these firms face strong incentives to divert government funds for their private interest. The government's efforts to combat public expenditure fraud face challenges: increased oversight can be costly or distortive, and bureaucrats themselves face weak incentives to combat waste and fraud as they face limited accountability mechanisms and are not personally responsible for the government's financial mismanagement.

Privatization is a potentially effective way to both elicit private information useful for the detection of fraud and also to provide incentives to catch fraudsters. The False Claims Act (FCA) is a federal law that allows whistleblowers to recover over-billed money for the government and receive a share of the recoveries. Uniquely, FCA whistleblowers conduct their own litigation on behalf of the government in federal civil court, combining the private information of whistleblowers with the private enforcement of law. This process has generated thousands of whistleblowing lawsuits and recovered tens of billions of dollars since the 1986 enactment of the law. In fiscal year 2018 alone, whistleblowers recovered $2.1 billion for the government from FCA lawsuits, for which whistleblowers were awarded $301 million (US Department of Justice, 2018). Despite the volume of lawsuits and recovered funds, there has been very little empirical evidence on the False Claims Act's effectiveness at both catching and deterring public expenditure fraud.

The issues of fraud in public expenditure are compounded in the provision of federal health insurance. The US federal government relies heavily on private firms to provide healthcare and reimburses these providers based on self-reported activity. Furthermore, much of healthcare is a credence good, meaning that neither patients nor non-specialist bureaucrats are able to effectively monitor doctors (Dulleck and Kerschbamer, 2006). This information asymmetry provides opportunities and incentives for healthcare firms to increase their profits through misreporting. In contrast, the federal insurer has a limited capacity for monitoring and enforcement. With the US government spending more than a trillion dollars per year on healthcare, even small shares of impropriety can be expensive, prompting concerns about the magnitude of healthcare fraud and generating interest in the ways to combat it. Correspondingly, 55% of False Claims Act whistleblower lawsuits are related to the federal healthcare programs.

In this paper, I examine the economics of whistleblowing under the False Claims Act and empirically measure the costs and benefits of private enforcement, with evidence from the Medicare program. Medicare provides an excellent setting to understand the effects of private enforcement on fraud more generally, as we can observe all relevant expenditures and many of the important social consequences of these policies. First, I model the decision of a whistleblower to litigate, as compared to socially optimal behavior, and discuss the key magnitudes needed to understand whether privatization is efficient. For my empirical analyses, I pair a novel dataset on whistleblower filings and their allegations with a large sample of Medicare claims data from 1999 to 2016 to measure the benefits and costs of whistleblowing. I estimate the deterrence effects of a set of the largest successful whistleblower cases, as well as the public costs of whistleblowing and its effects on patient health outcomes. Overall, I find large deterrence effects, small public costs, and no evidence of negative health effects on patients, indicating that private enforcement is an effective anti-fraud policy.

Private antifraud enforcement involves social benefits, private costs, and public costs, none of which are fully internalized by the whistleblower. In a simple model of whistleblowing litigation, I show how the decision of the whistleblower to litigate differs from the publicly efficient choice. By paying whistleblowers a share of the money recovered, the False Claims Act creates incentives for whistleblowers to conduct enforcement when they expect cases to be profitable, which is proportional to the public cost of the fraud already committed. Yet whistleblowing can

also have deterrence effects for which the whistleblower is uncompensated. Whistleblowing can change spending on the types of fraud the whistleblower identifies, called specific deterrence, and can also cause spillovers to inhibit fraudulent billing throughout the medical industry, known as general deterrence.[1] Private enforcement can also risk over-enforcement because whistleblowers do not bear the full costs of their litigation. Lawsuits have both public and private costs, including to the Department of Justice, the court system, and the attorneys of the plaintiffs and the defendants. Whistleblowing can also affect care decisions by providers, with either positive or negative consequences for patient health that are not internalized in the whistleblower's enforcement decision.

To measure these deterrence effects, I analyze the effects of whistleblowing on public spending in four case studies of enforcement related to the Medicare program. These four case studies reflect dozens of lawsuits against hundreds of defendants, grouped by similar allegations of fraudulent conduct. The cases selected are the highest dollar value antifraud enforcement for which I have Medicare data. I use a novel synthetic control methodology, which I call "staggered synthetic controls," to estimate counterfactual spending in the absence of whistleblowing. My method extends existing synthetic control methods by using donor control units that occur asynchronously and estimating a time shift for these units to align them with the treated unit. This extension allows the comparison between similar trends in spending that occur at different points in time, which improves the pre-period fit of the synthetic control group. In effect, this method extends the staggered difference-in-difference approach to synthetic controls.[2] I compare spending on types of medical care subject to a whistleblowing lawsuit against a synthetic control group constructed of similar types of care not affected by the lawsuit. The difference between these series measures the specific deterrence effect of the whistleblower lawsuits.

My results show that whistleblowing achieves a high level of specific deterrence, as measured by the change in spending on the procedures identified as fraudulent by the whistleblowers, netting out increases in spending on substitute procedures. I estimate that total specific deterrence effect of these four case studies is $18.92 billion in the first five years after the suits were filed. On average, specific deterrence is 6.8 times the case's settlement value, but with wide variation in this ratio across case studies. Importantly, these specific deterrence effects do not count the general deterrence value of these cases in deterring other types of fraud not identified by the whistleblower. Therefore, this estimate constitutes a conservative measure of the total deterrence caused by these suits. A variety of robustness checks support the high deterrence value found via the synthetic control method.

The public costs of whistleblowing are modest and are a small fraction of the estimated benefits. These costs include expenditures by the federal agencies that oversee and contribute to the litigation, including the courts, the Department of Health and Human Services, and the Department of Justice. Using data from federal budget reports, I estimate that total federal expenditures were less than $108.5 million in 2018. This indicates that whistleblowing has an extremely high return on investment for federal resources, and that privatization is a promising mechanism for antifraud enforcement.

Finally, I examine the effects of whistleblowing on provider care decisions. Whistleblowing has the potential to change provider care decisions by changing the compliance requirements, litigation risk, and profitability of care that doctors conduct. The net impact of whistleblowing on patient health is ambiguous *a priori* and depends upon how whistleblowing deters or encourages different forms of care. To examine this question, I conduct a case study

---

[1]The distinction between "specific" and "general" enforcement and deterrence comes from Shavell (1991).

[2]Unlike difference-in-difference designs, but similar to other synthetic control applications, my analysis has a single treated unit in each case study and multiple untreated controls. Recent papers have begun to apply synthetic control weighting methods to difference-in-difference designs, focused on the case of multiple treated units– see, e.g., Ben-Michael, Feller, and Rothstein (2019). Those papers apply the weighting procedure of synthetic controls to the difference-in-difference setting; this paper instead applies the asynchronous method of staggered difference-in-difference onto a synthetic control setting.

on kyphoplasty, a spine procedure for patients with vertebral compression fractures from osteoporosis. Kyphoplasty is chosen because it is the largest case study I conduct to study financial deterrence for which I expect any change to patient health – (by contrast, one larger case study involves accounting manipulation and should have minimal patient health effects). A set of whistleblower lawsuits alleged that many hospitals fraudulently admitted patients for inpatient kyphoplasty rather than perform this procedure outpatient. Kyphoplasty is linked to decreased mortality in the medical literature, and there was a reduction in inpatient kyphoplasty following the whistleblowing lawsuits, motivating a concern about inadvertent patient consequences. I model the effects of kyphoplasty on Medicare patient death among two cohorts of roughly 8 million patients each, from before and after the lawsuits. My model interacts patient covariates and medical claims history with treatment and produces predcited treatment effects of kyphoplasty for each patient. These effects provide a scale by which to measure how beneficial it is for a patient to receive treatment.

I find that in the case of kyphoplasty, whistleblowing had positive effects on patient care overall. Following whistleblowing, there was better targeting of the procedures to those patients predicted to benefit. Patients for whom it is expected that kyphoplasty increased mortality were 7% less likely to be treated, while patients with reduced mortality if treated were 7% more likely to be treated. This targeting change was concurrent with a substitution from inpatient kyphoplasty to less expensive outpatient kyphoplasty and vertebroplasty, a close substitute. This indicates that whistleblowing can have positive effects on care delivery by changing the incentives in the care decision process by providers, even while reducing spending.

While there has been substantial disagreement in the public sphere over the value of the False Claims Act, there has been little empirical evidence on the effects of this law.[3] Engstrom (2012; 2013) presents descriptive statistics on FCA cases and settlements but does not measure the law's effects on providers and spending. In the accounting literature, Heese (2018) shows that hospitals prosecuted under the FCA are less likely to participate in broad measures of overbilling. In the health literature, Howard and Desai (2020) show that FCA investigations lower angioplasty volume in investigated hospitals, and Howard and McCarthy (2021) show that False Claims Act enforcement deters the overuse of implantable cardiac devices. No paper has quantified the financial deterrence effects of whistleblowing across a generalizable sample of different cases, nor addressed the efficiency of the Act.

This paper also relates to a broader literature on private enforcement and deterrence. Becker and Stigler (1974) suggest the privatization and marketization of enforcement as a way to align the incentivizes of enforcers with those who benefit from the enforcement. Landes and Posner (1975) formalize the theory of private enforcement, and Polinsky (1980) compares public and private enforcement for the imposition of fines. Shavell (1991) formulates the differences between specific and general enforcement of law used here. While the legal literature on private enforcement theory is robust, there has been limited empirical evidence of private enforcement in practice due to data limitations and the relative rarity of private enforcement mechanisms. Notably, private enforcement is widely used to combat self-dealing in securities law; Djankov et al. (2008) and Jackson and Roe (2009) discuss the efficacy of this form of private enforcement with mixed results.

Finally, this paper relates to the economics literature on healthcare fraud more broadly, which has largely not

---

[3]Proponents of whistleblowing point to the volume of settled cases and the billions of dollars recovered as evidence of the effectivity of the Act. Attorney General Eric Holder said in a 2012 press release: "In the last quarter century, the False Claims Act's success has been unparalleled with more than $30 billion dollars recovered...and $8.8 billion since January 2009" (United States Department of Justice, 2012). Detractors suggest that profit-seeking whistleblowers use civil litigation to force settlements from providers, regardless of the validity of their allegations. In an *amicus* brief to the Supreme Court, the US Chamber of Commerce, a pro-business organization, wrote: "[whistleblowers] can extract settlements from defendants averse to high discovery costs, the risk of large losses, and...reputational harms" (Chamber of Commerce, 2015). Kwok (2013) uses data on whistleblowers and plaintiffs attorneys to show that spurious cases and filing mills are not an issue in FCA litigation.

discussed whistleblowing. Silverman and Skinner (2004) and Dafny (2005) describe the financial incentives for mis-reporting (particularly upcoding) among hospitals. The types of fraud described in those papers were ultimately litigated by whistleblowers under the False Claims Act; this paper fills a gap in the existing literature by evaluating the whistleblowing program that seeks to catch and inhibit the fraud described in those studies. Becker et al. (2005) show that increased state-level Medicaid anti-fraud enforcement by government investigators lead to declines in fraudulent treatment, with no negative patient health outcomes. In recent work, Fang and Gong (2017) use a measure of the time spent on procedures to detect providers who bill Medicare for too many hours. My analysis of the health consequences of whistleblowing also reflects the analysis of malpractice litigation in Kessler and McClellan (1996), who similarly find no observable harm to patients.

This paper is organized as follows. Section 2 describes the institutional details of the False Claims Act, and Section 3 models the economics of private enforcement. Section 4 describes the data and provides stylized facts about FCA lawsuits and recoveries. Section 5 measures specific deterrence on a set of case studies using the staggered synthetic control methodology. Section 6 addresses the costs of private enforcement. Section 7 presents the effects of the kyphoplasty lawsuits on patient care, and Section 8 concludes.

2 Background

Medical care has a fundamental information asymmetry among providers, insurers, and patients (Arrow, 1963), which creates opportunities for misreporting. Patients are less informed than doctors about the care they need, making health-care a credence good: patients are not sufficiently able to monitor doctors for bad behavior (Dulleck and Kerschbamer, 2006). Conversely, insurers have limited means of directly observing medical conditions or treatments, and rely on provider's claims for payment. This information asymmetry provides opportunities for misreporting by providers, whose billing practices tie directly to their profits. It is difficult to uncover misreporting using top-down enforcement, as insurers often lack other sources of information besides the provider's claim and supporting documentation, which can be manipulated.

When the insurer is the federal government, as is the case with Medicare and Medicaid patients, these problems are exacerbated. Medicare and Medicaid are massive programs, spending respectively around $700 and $400 billion per year (Congressional Budget Office, 2019), creating bureaucratic issues due to the sheer volume of claims. Indeed, the Government Accountability Office (GAO) estimates that $48 billion (8%) of Fee-for-Service Medicare expenditures in 2017 were "improper," that is, they lack necessary documentation to ensure the correct amount was paid to the right person for a valid claim (United States Government Accountability Office, 2019). Not all improper payments are fraudulent, and not all fraud is captured by improper payment measures, but this figure underscores the opportunism that may arise from expensive and overwhelmed federal programs.

With these issues in mind, in 1986 Congress amended the False Claims Act to enable whistleblowers to directly conduct lawsuits against those who overbill the government (United States Department of Justice, 2012).[4] The False Claims Act applies to all claims for payment made to the federal government, but has largely been used against healthcare fraud, overbilling, and misreporting. Under the False Claims Act, individuals who uncover misreporting against the US government, themselves often healthcare workers (e.g. a hospital employee), hire their own attorneys and sue those filing false claims in federal civil court.[5] The whistleblower sues *qui tam*, i.e., on behalf of the US

---

[4]The FCA was amended in 1986, but originally existed during the Civil War to combat fraud against the Union army. It was ineffective and out of use in the 20th century before the 1986 amendments.

[5]Critics may assert that whistleblowers could extort payments from firms without filing a False Claims Act lawsuit, using private nondisclosure

government. These civil court cases have 3 parties: the whistleblower, the defendant, and the US government.[6] In some cases, the Department of Justice intervenes in what it believes to be a lucrative lawsuit by assigning its own attorneys, conducting additional investigation and overseeing litigation.[7] In other cases, the whistleblower does not receive federal support, and either pursues the case alone or drops it. All cases are filed under seal, meaning the defendant is not immediately notified of the filing, giving the government an opportunity to investigate and elect to intervene before the defendant is made aware. The Department of Justice must approve any settlements between the whistleblower and the defendant, regardless of their intervention status.

False Claims Act lawsuits can be high stakes for all parties involved. These cases are conducted in civil court, and the burden of proof is the preponderance of the evidence, i.e. "more likely than not." Violators of the False Claims Act are jointly and severally liable, and so each defendant to an FCA case involving multiple parties could be held responsible for the full damages (O'Neil et al., 1995). Because litigation is expensive, few cases go to trial; unsuccessful cases are often voluntarily dismissed by the whistleblower, and clear-cut cases are settled. In successful cases, the federal government can recover up to 3 times the amount of the proven false claims from the defendants, plus potentially large criminal fines. Upon settlement, the whistleblower is entitled to 15–25% of the recovery amount if the government intervened, and 25–30% if the government did not intervene. The government and whistleblower only recover funds from the lawsuit if the defendant settles or the judge rules against them. Whistleblowers regularly earn 6-figure payouts and above from these cases, of which their attorneys, working on contingency, take around 30%. Defendants are also often issued criminal fines and can be sued for legal fees by successful whistleblowers; furthermore, in egregious cases the Department of Justice can pursue criminal action against responsible individuals. Civil enforcement is compounded by the use of Corporate Integrity Agreements, where defendants who settle agree to additional federal oversight, or by the use of exclusion of the provider from the Medicare and Medicaid programs. Section 4.1 provides summary statistics on FCA lawsuit outcomes.

The ability for the whistleblower to conduct the lawsuit in lieu of the government creates a profit motive for rooting out impropriety that may be otherwise lacking in the federally administered programs. This profit motive is in contrast to the usual incentives of federal bureaucrats, and thus can alleviate principal-agent problems within the government that can cause inefficient investment in monitoring and enforcement. Prosecution initiated by the government also has capacity constraints due to the limited resources of the Department of Justice, while privatized enforcement creates a market for whistleblowing information and generates substantially more litigation than the federal government conducts alone.[8]

Unlike other whistleblower programs, such as the IRS or SEC whistleblower programs, False Claims Act prosecution is conducted directly by the whistleblower. This eliminates the prosecutorial discretion component of government

---

agreements. From an institutional standpoint, this is unlikely because privately contracting with a whistleblower does not release the firm from liability for similar claims from other whistleblowers or from direct government enforcement, while such release of liability is standard in an FCA settlement. However, firms do face pressure to self-disclose billing errors to Medicare administrators, as this can circumvent the FCA process and allow an out-of-court repayment without large punitive damages.

[6]In legal terms, the whistleblower is called a "relator" to the civil lawsuit.

[7]The Department of Justice retains the option to intervene in cases with the intention of dropping them, in order to facilitate the dismissal of cases it believes to lack merit. During the period of study in this paper, from 1986 through 2012, this was not practiced. However, an internal 2018 Department of Justice memo (the "Granston Memo") promoted this type of intervention. This new policy has led to a variety of ongoing litigation, and the long term policy of the Department of Justice remains unclear. Latham and Watkins White Collar Defense and Investigations Practice (2020)

[8]The False Claims Act has a separate provision for direct enforcement by the Department of Justice without whistleblowers, if for some reason the government has information about misreporting or fraud against federal programs without a whistleblower filing a lawsuit. That portion of the False Claims Act is not studied in this paper. Since 1993, FCA lawsuits filed by whistleblowers have exceeded FCA lawsuits by the government; in 2016, there were 501 new whistleblower suits to 69 federally initiated suits (US Department of Justice, 2018). These statistics are only included as a point of comparison.

enforcement, and may lead to the litigation of cases for which there is little social harm or even explicit misconduct. Because there is no restriction on the filing of False Claim Act lawsuits, there is a potential for frivolous litigation by profit-motivated whistleblowers seeking a settlement. These problems are exacerbated by institutional details that have been analyzed as contributing to frivolous lawsuits in other legal literature: whistleblowers retain attorneys on contingency, face a potential for a very high-value settlement, and may take advantage of fee-shifting provisions in the case of successful lawsuits (see Bebchuk and Chang (1996) and Bebchuk and Kelement (1998) for a theoretical discussion of frivolous suits). Defending against FCA lawsuits can be expensive, and defendants should settle if the expected cost of settlement is below the expected costs of fighting the lawsuit (and potentially losing), regardless of the truthfulness of the whistleblower's claims.

There are some institutional barriers to whistleblowing that deter low-quality cases. First, whistleblowers are not allowed to represent themselves in court (United States District Court, D.C., 2003). Due to the costs of litigation, and the fact that plaintiffs' attorneys work on contingency, plaintiffs' attorneys have incentives not to take on low-quality lawsuits. This provides a barrier to filing frivolous cases. Furthermore, FCA cases are most likely to be successful if the government intervenes, due to the resources and investigatory power the federal government brings when litigating a case. Since low-quality lawsuits are unlikely to generate an intervention from the federal government, this further exacerbates the unwillingness of plaintiffs' attorneys to self-fund any low-quality cases. Empirically, Kwok (2013) studies data on FCA whistleblower attorneys and finds no evidence for "filing mills", i.e. law firms pursuing a large volume of low-quality cases.

3   Theory of Private Antifraud Enforcement by Whistleblowers

3.1   Model

Consider the following simplified model of private enforcement against Medicare fraud under the False Claims Act. This model is motivated by other general models of private litigation not specific to the FCA, including Shavell (1982) and Spier (2007). This model explores the divergence between whistleblower's private incentives to litigate and the decision a social planner would take.

A whistleblower receives a private signal about fraud occurring against the government, in the form of the probability π that they can win a False Claims Act lawsuit, and the amount of fraud being committed as measured by the expected federal recovery amount R(including fines). Whistleblowers are not permitted to represent themselves (United States District Court, D.C., 2003), and the whistleblower makes a joint decision with their attorney of whether to file suit; both parties must agree. The whistleblower and their attorney observe private costs C_w associated with conducting litigation. Attorneys working on contingency expend resources pursuing litigation, most notably their time, which carries a high opportunity cost. A whistleblower attorney may be able to recover some of the costs of litigation from the defense in the case of a successful suit, although this fee-shifting is not always practiced. If the case is successfully pursued by the whistleblower-attorney pair, they expect to receive a share s of the recovery. Then, a risk neutral whistleblower-attorney pair would choose to litigate if:

$$\pi s R > C_w \tag{1}$$

The expected whistleblower compensation $\pi s R$ creates incentives for the whistleblower to come forward with their private information about fraud or misconduct, and can alleviate personal and professional costs arising from

whistleblowing on one's employer, as well as the realized costs of litigation, all captured in C_w.

The whistleblower and attorney's decision reflects the fact the decision to litigate is increasing in the probability of settlement as well as the magnitude of the recovery, and decreasing in the expected litigation costs. This indicates that, all else equal, whistleblowers and their attorneys are more likely to litigate against large frauds, as well as frauds for which they have the greatest evidence, as evidence can both increase the likelihood of a successful suit π and also decrease investigatory litigation costs.

Consider instead the decision of a social planner of whether a particular fraud should be litigated against. Whistleblowers fail to internalize the costs of litigation that they do not bear. The defendant must spend C_d to defend against the lawsuit, and the lawsuit also has public cost C_p, reflecting the resources expended by the federal government in oversight and trial of the case.

Litigation by the whistleblower can also produce benefits that whistleblowers fail to internalize. Whistleblowing produces specific deterrence effects D_s, measured in cost savings by the government for the type of fraud the whistleblower deters. Figure 1 shows the relationship between spending, damages and specific deterrence. I define specific deterrence as the difference between spending with and without whistleblowing, integrated after the time of the lawsuit:

$$D_s := \int_{t_s} (S_f - S_w)\delta^{t-t_s} dt \tag{2}$$

where t_s denotes the time the lawsuit initiates, S_w is spending with whistleblowing, S_f is spending with unaddressed fraud, and δ is a discount factor. In a circumstance where whistleblowers cause a decrease in one specific type of spending but an increase in a substitute procedure, $S_f - S_w$ reflects the net changes in spending.

In contrast, the recovery R is proportional to the damages the government faces. As shown in Figure 1, damages are the integral of fraudulent minus non-fraudulent spending up to the lawsuit:

$$R \propto Damages = \int^{t_s} (S_f - S_n)\delta^t dt \tag{3}$$

where S_f is fraudulent spending without whistleblowing and S_n is spending without fraud. Spending with whistleblowing S_w and spending without fraud S_n may be the same if whistleblowing completely deters fraud and returns spending to a non-fraudulent level. As shown in Figure 1, the relationship between the magnitude of damages and of specific deterrence are governed by the amount of fraudulent spending as well as the extent to which whistleblowing curbs fraudulent behavior.

In addition to specific deterrence, whistleblowing can cause general deterrence D_g, not shown in Figure 1. Whistleblowing against a particular fraud may cause providers committing unrelated frauds to stop doing so because they fear being caught. Moreover, the existence of whistleblower lawsuits may deter providers from committing other frauds in the first place. The magnitude of general deterrence is difficult to measure empirically, as it concerns behavior not identified by the whistleblower. However, general deterrence effects must decrease spending on other fraudulent or misreported procedures, because providers are less likely to commit fraud given increased scrutiny. Therefore, I assume $D_g > 0$.

In addition to the financial costs and benefits discussed above, whistleblowing cases may also have impacts on provider decisions over patient care. False Claims Act cases may inform providers' care decisions as they seek to comply with the shifting landscape of regulation and litigation risk. These changes in provider behavior can be consequential to patient health outcomes. These changes to patient health outcomes may pose an additional cost H_c if whistleblowing distorts care away from the social optimum, or may provide additional benefits H_b if whistleblowing

7

corrects fraudulent behavior that jeopardizes patient health. This relates to similar research on malpractice liability, in particular Kessler and McClellan (2002), whose model predicts that malpractice liability can cause either inefficiently high or low levels of care. I expect that the health effects of whistleblowing may differ between lawsuits, with some providing positive changes, and others causing negative changes.

From a social welfare perspective, private enforcement by the whistleblower is efficient if the total expected benefits exceed the total costs of the lawsuit:

$$\beta[\pi(1-s)R + \pi D_s + D_g] + H_b > C_w + C_d + C_p + H_c \tag{4}$$

The probability of settlement π gives the probability that the government gains the recovery $(1-s)R$, and also governs the probability that specific deterrence effects D_s are produced. I define specific deterrence as the effect of a successful whistleblower lawsuit. General deterrence effects D_g, measured in spillovers of this case onto other fraud decisions, can be produced even if the whistleblower's case is not successful; however, the magnitude of general deterrence may vary with factors that correlate to the probability the case is successful, such as the quality of evidence or the magnitude of alleged fraud. Costs, as well as changes to healthcare outcomes associated with litigation, are borne regardless of the outcome of the suit. The whistleblower's share sR is a transfer from the defendant to the whistleblower and does not factor into social welfare.

The social planner values the government's portion of the recovery $\pi(1-s)R$ and the deterrence effects $\pi D_s + D_g$ with a multiplier $\beta > 0$. This reflects the fact that the disgorgement of previously misappropriated funds is a transfer from the defendant to the government, and the deterrence effects are cost savings to the Medicare program. Therefore, the social benefit of these values is the marginal benefit of public funds. Exact values for the coefficient β depend on how disgorged funds are used. FCA lawsuit recoveries are remitted to the Medicare trust fund, and deterrence values are dollars not spent by the Medicare trust fund. If these additional funds relax the federal budget constraint and reduce taxation, β may be greater than 1, as they remove the distortionary effect of taxes raised elsewhere. If the taxes increase the total expenditure on healthcare, β should reflect the marginal social value of Medicare expenditure. Hendren and Sprung-Keyser (2020) provides a deeper discussion of the marginal value of public funds (MVPF) and the welfare differences between decreased taxation and increased spending. They estimate the MVPF for many adult health policies between 0.5 and 2, and estimate that the MVPF of the introduction of Medicare was 1.63.

The social value of whistleblowing also depends on how the social planner values the disutility of the defendant whose fraudulent funds were disgorged. This paper does not make the structural assumptions necessary to measure changes in defendant utility, nor do I estimate how much the social planner should weight the defendant's disutility of repaying funds they had ostensibly misappropriated. These parameters can also be captured by the value of the coefficient β.

3.2 Discussion

The difference between the whistleblower's decision as captured in Equation 1 and the social planner's choice in Equation 4 highlight some of the key issues facing private antifraud enforcement.

Whistleblower cases have the potential for valuable specific deterrence effects D_s, which contribute to the public good but not to the whistleblower's decision. Following a lawsuit, both the defendants and other providers of the same care face incentives to change their behavior to avoid further litigation or to comply with the terms of their settlement agreements. Defendants face potential exclusion from the Medicare and Medicaid programs for noncompliance with

8

their settlements. Because the defendants may be only a small share of those committing litigable behavior, these specific deterrence effects have the potential to affect providers far exceeding the scope of the settlement.

One might expect that providers who commit "rational fraud" do so having fully internalized the expected costs of their behavior, and observing settlements would not affect their decisions. However, observing settlements can either update other providers' beliefs about being caught, or increase the salience of the expected costs, thus causing behavioral changes and specific deterrence effects. In addition, behaviors that constitute litigable FCA violations may be "gray areas" of billing or care, in which case settlements can draw a clear line on what is acceptable behavior, and can prompt rule changes and clarifications from the Medicare administrators.

This model gives rise to a set of cases where the private and public interests diverge. If damages prior to the whistleblower's signal are small, the expected recovery R may be too small to be worth pursuing by the whistleblower, and the inequality in Equation 1 fails to hold. Yet from a social welfare perspective, these cases might be valuable to litigate if the deterrence value would be large. For example, a small trickle of fraud in perpetuity could have a low recovery but a high deterrence effect if enforced against. In contrast, there are potential circumstances in which the specific deterrence values are small. Specific deterrence is the difference between spending with and without whistleblowing, and when these values are similar then specific deterrence is small. This could occur when the increase in spending due to fraud all occurs before the whistleblower files, and future spending would look the same with or without whistleblowing. In this circumstance, the settlement serves as a transfer from the defendant to the government and whistleblower for past bad actions, but there is no specific deterrence. However, there may be general deterrence effects, if observing this transfer changes others' beliefs about their own enforcement probability or about the profitability of fraud. Another circumstance with little specific deterrence effect is one in which whistleblowing is not meaningful; for example, if fraud continues to be profitable even following a settlement, whistleblowing may not deter future bad behavior of the same sort, and $S_w - S_n$ is small. In these circumstances, whistleblowing is potentially inefficient because the settlement only serves to correct retrospective damages, and the lawsuit incurs its full costs without providing social value into the future.

The value of the deterrence effects $D_s + D_g$ is policy-relevant in evaluating the compensation of whistleblowers. Whistleblowers are paid a portion of the settlement recovery, which is itself proportional to the amount of damages due to pre-settlement overbilling. Therefore, whistleblowing compensation is purely *retrospective*. However, the value of whistleblowing depends on both the settlement and the deterrence effects, the latter of which does not factor into whistleblower compensation. This disconnect between whistleblower compensation and whistleblower value-added may indicate that whistleblowers are inefficiently compensated, as they do not the public good element of the deterrence they provide. As an alternative policy, whistleblowers could be compensated based on both settlement and their ex-post deterrence effects, for example through a contract that pays the whistleblower for a proportion of the deterrence realized after their suit. This discussion ties directly to the literature on incentives for private enforcers, as discussed in Polinsky (1980).

The timing of whistleblowing also factors into its social benefits as well as the whistleblower's compensation. The faster that fraud is litigated against, the smaller the retrospective damages and, therefore, the smaller the whistleblower's share sR. This could in theory cause whistleblowers to increase their payout by waiting before filing their lawsuit, if fraud is ongoing, to allow damages to realize. However, these effects are mitigated by a priority race, in which the first-to-file whistleblower generally receives the bulk of the compensation. The False Claims Act also has a statute of limitations of at most 10 years from the date of the fraud to the filing of the whistleblower lawsuit (31 US Code Section 3731, 1986). From a social welfare perspective, the timing of the whistleblower lawsuit is ambiguous,

9

because smaller recoveries due to earlier litigation are reflected in greater deterrence amount. In practice, plaintiff attorneys report that they tend to file the lawsuit as quickly as they are able to put together a good case.

The efficiency of this private enforcement regime relies on the extent to which the benefits of whistleblowing outweigh its costs. This motivates an analysis of both the deterrence effects and health consequences of False Claims Act lawsuits. Section 5 undertakes an exercise to measure the deterrence effects of the largest whistleblower cases, and Section 7 undertakes a case study of the health effects of a set of whistleblower lawsuits against a spine surgery procedure.

4   Data and Descriptive Statistics

The data for this project come from a variety of complementary sources which aggregate information on whistleblower cases and their downstream impacts on medical care provision and patient health outcomes.

Data on whistleblowing at the lawsuit level comes from a FOIA request I conducted on the Department of Justice in 2018 for data on all completed (settled or dismissed) whistleblower-filed FCA suits (Department of Justice, Civil Division, 2018).[9] These data describe almost 6,000 whistleblowing cases and include information on the defendant, whistleblower, filing date, federal agency to which the case relates, federal court district of filing, government intervention election status and date, settlement amount, and whistleblower share. These data start with the introduction of the law in 1987, and the coverage declines after 2012, as many newer cases are still under seal. These data are used for descriptive statistics and stylized facts in section 4.1, as well as for providing supplementary information on whistleblower lawsuits for each case study in Section 5. Appendix A describes the data cleaning process.

The FOIA data from the Department of Justice do not contain allegations of conduct by the defendants, which is necessary to trace the effects of lawsuits in the Medicare claims data. To find such allegations of fraud, I scraped the Department of Justice "Justice News" archive website for all press releases related to Medicare and whistleblowing. For the cases analyzed in this paper, I also collected whistleblowers' original court filing documents (complaints), settlement agreements, and other court documents from a variety of sources. These documents detail exact filing dates, settlement timing, allegations of fraud, and the conduct covered by the settlement agreements. Sources for these documents include the federal court record system (PACER), the Department of Justice digital archives, SEC filings of publicly traded companies, and the legal database of Taxpayers Against Fraud, a not-for-profit supporting whistleblowers' attorneys. Combined with the press release and FOIA data, the court filings give a complete picture of the allegations and outcomes for a subset of the whistleblower lawsuits for which I conduct case studies. Section 5 describes this process and presents its results.

Data on Medicare claims and payment are necessary for the analysis of the medical and fiscal impacts of whistleblowing cases. My available data include 100% samples of Fee-for-Service Medicare, i.e. Parts A and B, from 1999–2016, for inpatient, outpatient, hospice care, durable medical equipment, home health care, and skilled nursing facilities; and 20% samples of the "carrier" files that reflect physician office visits (Center for Medicare and Medicaid Services, 1999-2016). These data, containing mostly 100% samples of each type of care over nearly 20 years, cover tens of billions of claims from hundreds of millions of patients. Section 5 presents the methodology by which I selected whistleblowing cases for analysis, which translates into the usage of these data. Medicare data are used only as they related to each case presented there, and for the analysis of patient health outcomes in Section 7. As such,

---

[9]This data set is similar to that used by Engstrom (2012; 2013), which also came from a DOJ FOIA request. However, to access the most recent data available, I conducted an original FOIA request.

only a portion of the available data is used in these analyses, reflecting the "needle in a haystack" aspect of Medicare overbilling.

Information on the costs of whistleblowing comes from a different set of data. The Department of Justice does not publish data on the costs of FCA lawsuits directly. Data on public expenditures related to civil enforcement were culled from federal budget reports, particularly the 2018 Health Care Fraud and Abuse Control Program Annual Report from the Department of Health and Human Services and the Department of Justice (The Department of Health and Human Services And The Department of Justice, 2019). Other data on costs were gathered from the Department of Justice Qui Tam Fraud Statistics (US Department of Justice, 2018) and the budget of the federal judiciary (Administrative Office of the US Courts, 2017).

4.1 Descriptive Statistics about False Claims Act Lawsuits

Table 1 presents descriptive statistics about False Claims Act lawsuits. I observe 5,967 lawsuits, of which 3,269 (55%) are healthcare-related. Of the healthcare cases, only 35.7% result in a recovery of funds; the rest were dismissed by the whistleblower, the judge, or the Department of Justice. This points to a high level of cases for which the federal government receives no compensation, underscoring questions about the efficiency of the False Claims Act. Government intervention is nearly synonymous with a successful case: while the government only intervenes in 32% of lawsuits, 91% of such cases settle. This reflects both the selection effect, i.e. that the government chooses promising lawsuits to intervene in, as well as the treatment effect of assigning government attorneys to conduct litigation.

Both the settlement amounts and the whistleblower shares have high variance reflecting very long right tails. The median settlement amount for healthcare-related whistleblower lawsuits is $1.5 million, but the mean is $22.7 million and the standard deviation is $87 million. Total settlements amount to $26.47 billion, for which whistleblowers were paid a total of $4.23 billion, with a median of $250,000 per case. Similarly, whistleblower cases take a highly variable amount of time. The median healthcare whistleblower lawsuit length, defined as the time from filing to case closure, is 964 days (2.6 years) with a standard deviation of 800 days (2.2 years).

Appendix Figure A1 shows the trend of healthcare whistleblowing cases by year of filing and whether they end in a settlement. Settlements rose between 1990 and 1995 to around 75 cases per year, and have stayed rather constant since. Conversely, the total number of cases and the share of dismissed cases have both risen substantially since 1987, and continue to grow. Cases that are ultimately dismissed now constitute the majority share of whistleblowing. Naturally, the settling or dismissal of cases does not reveal their underlying merit – some meritorious cases are dismissed (for example, due to cost reasons) while some frivolous cases settle, (for example if the defendant is particularly risk-averse).

While the number of settled cases has remained steady since 1995, total settlement dollars have risen immensely. Total settlements were just $80.6 million in 1995, when the number of settled cases reached its steady state of around 75 per year. However, settlement totals exceed $4 billion in 2012, the last year of the data. The 2012 total was in a large part due to a single $1.5 billion settlement against GlaxoSmithKline for allegedly promoting its pharmaceuticals for non-FDA-approved uses.

The Department of Justice Data also include lawsuits from outside of the medical field, and exhibit the broad use of the False Claims Act. Medical-related suits, those categorized by the DOJ as relating to the Department of Health and Human Services, the Food and Drug Administration, or the Center for Medicare and Medicaid Services, constitute 55% of cases. Suits regarding the Department of Defense account for 11% of the nearly 6,000 whistleblower lawsuits, and cases have arisen from nearly all parts of the federal government, including the Department of Education (3%

of cases) and the Goods and Services Administration (2% of cases). The use of FCA whistleblowing outside of the medical field is beyond the scope of this paper and presents an opportunity for future research.

5 Deterrence Effects: Method and Results

The deterrence effects of whistleblower lawsuits are a key component in the economic tradeoffs described in the model in Section 3. If deterrence effects are large, whistleblower lawsuits not only provide recovered funds for the government, but also save the government money in the form of fraud not committed.

As discussed in Section 3, deterrence from a whistleblowing lawsuit takes two forms: specific deterrence, from changes in spending on the type of care that the whistleblower identifies as fraudulent, and general deterrence, from reduced spending on other types of fraud due to increased litigation risk. In simpler terms, specific deterrence is the main effect of the lawsuit, and general deterrence captures the spillover onto other types of fraud. This analysis measures the dollar value of the specific deterrence of whistleblower cases.

The section proceeds as follows: first, I describe the econometric method for the analysis, which uses staggered synthetic controls to estimate the treatment effect of whistleblower lawsuits on healthcare spending. Second, I describe the lawsuits that serve as case studies, which are the largest whistleblower cases for which I have data. Finally, I apply the method to these case studies and find that whistleblowing produces large specific deterrence effects, saving the government nearly $19 billion in deterred fraud from just $1.9 billion of settlements. The cases I analyze account for roughly 7% of total whistleblower healthcare settlements. I conduct a series of robustness checks that support these results. Importantly, specific deterrence does not count the spillovers of these lawsuits onto other types of fraud, and therefore this estimate provides a lower bound of the total deterrence effects of the FCA.

5.1 Method

The measurement of deterrence requires an analysis of a counterfactual, between the real world in which enforcement happened and one in which it did not. Synthetic controls, first introduced in Abadie and Gardeazabal (2003), provide a method by which to produce such a counterfactual. Here, the outcome of interest is spending on the type of medical care treated by whistleblowing, and the treatment effect of interest is the change in spending following whistleblowing. The control groups are other types of medical care that are not treated by whistleblowing, but saw similar rises in spending. These increases may have been due to unchecked fraud or other profit-centered changes in billing practice. Synthetic controls use untreated control groups to construct a series that most closely matches the treated unit in the pre-treatment periods. The difference between the treated unit and the synthetic control group in the post-treatment periods estimates the treatment effects.

Traditional synthetic controls make the assumption that the counterfactual of the treated unit in the absence of treatment can be captured by using a set of controls with contemporaneous time patterns (for example, in the factor model given by Abadie et al. (2010)). However, whistleblowing often affects types of care with unusual trends: they exhibit high growth in spending and claims, potentially driven by the improper conduct of the defendants. This means that there may be few control units on similar trajectories, and a sparse donor pool can lead to a poor fit in the pre-treatment periods when the contemporaneous control groups do not have such trends.

Rather than comparing the treated unit to contemporaneous controls, I consider control groups that saw a similar rise in spending at other points in time but were not subject to enforcement by whistleblowers within the sample

period. The rise in spending on control units used here were also potentially driven by fraud; the specific control units given weight by the method are presented in the results section, and many are indeed other forms of fraud that were eventually enforced against, much later. This method makes the assumption of a similar growth trajectory at different points in time between treated and untreated units in the absence of whistleblowing. Appendix B.1 presents an econometric model consistent with this assumption that motivates the use of staggered synthetic controls, modeled after Abadie et al. (2010). Under this model, the staggered synthetic control method can estimate the counterfactual spending of the treated unit, as though it were untreated, by producing a synthetic control group constructed as a weighted combination of untreated units.

This modification of synthetic controls relies on different assumptions than standard synthetic controls. In effect, this model assumes that the "life cycle" of the fraudulent procedure dominates any common time effects. However, this raises concerns about time-varying confounders, as well as concerns about spillovers between treated units and control units due to their asynchronous timing. Section 5.5 presents a series of robustness checks that further address these issues, which find similar estimates of the deterrence value.

Figure 2 provides a simple graphical explanation of the staggered synthetic control method for two controls, one shifted forward in time and one shifted backward. The method is implemented with a two step procedure. First, I estimate a time shift for each control unit by finding the time shift with the best pre-period fit between the control group and the treated unit. Second, I construct a synthetic control group by estimating weights for the time-shifted controls. The resulting synthetic control group is used to estimate the counterfactual of the treated unit in the absence of enforcement.

For each case study, I estimate the deterrence effect by integrating the difference between the observed spending and the synthetic control in the post-lawsuit period. Appendix B.2 provides the technical details of the implementation of the shifting, weighting, and deterrence integral estimation. When the type of care treated by whistleblowing has a clear substitute, e.g. inpatient and outpatient medical care, I consider the net change in spending by applying the synthetic control method to both the treated unit and its substitute.

The deterrence values presented in this paper are computed using 5 years of post-treatment effects at a 10% discount rate, which is a conservative estimate. In the absence of whistleblowing, fraud or abuse may have continued indefinitely into the future, in which case the total deterrence effect would be a perpetuity, providing value at all later periods. Rather than assume that the deterrence effects persist indefinitely, the use of 5 years of post-treatment effects avoids excess extrapolation. Fundamentally, the specific deterrence being measured here is a lower bound of the total deterrence effect, as it also does not include the general deterrence effect, i.e. the spillover of these lawsuits onto deterring other types of fraud.

This method extends the use cases for synthetic controls. Traditional synthetic controls are only useful when there are contemporaneous control units that experience similar pattens to the treated unit in the pre-treatment period. This fails in circumstances where the treated unit is on a rarely seen trajectory. In these circumstances, the staggered synthetic control method can estimate the treatment effect on the treated unit from the pattern of the other similar, untreated units that occur asynchronously. In this sense, it extends the staggered difference-in-difference approach to synthetic controls.[10] This method could be used for a variety of applications in circumstances where traditional synthetic controls produce a poor pre-period fit but the researcher would like to use control units from different points in time.

In order to conduct inference on my results, I employ a permutation test as per Abadie et al. (2010). Each synthetic

---

[10]See e.g. Athey and Imbens (2018) for a discussion of the econometrics of staggered difference-in-difference.

control is substituted in for the treated unit, and the same two-step procedure detailed above is performed, fitting leads and lags and constructing weights, using all other controls. These weights give a synthetic control unit for the placebo, from which the deterrence measurement can be computed. The deterrence effects corresponding to each control unit form an empirical distribution against which the deterrence effect of the treated unit can be compared.

5.2 Case Selection

An ideal study of whistleblowing would analyze every whistleblower lawsuit to compute deterrence effects; however, such an analysis is impossible. There is no complete database that provides information on the allegations of fraud of whistleblower lawsuits, including the Department of Justice FOIA Data. Each lawsuit must be researched individually through Department of Justice press releases and publicly available federal court filings. Some whistleblower documents remain under seal, and so the whistleblower's allegations are not available whatsoever. Even for unsealed lawsuits, finding corresponding claims in the Medicare data requires an extensive understanding of medical billing for the procedure, which has changed extensively over time.

The difficulty of mapping from lawsuits into the medical claims data motivates a case study design. The summary statistics presented in Section 4.1 show that a few large whistleblower cases dominate the settlement totals and amounts paid to whistleblowers. Therefore, I undertake a case-study based design of the largest whistleblower cases for which I have data.

Many lawsuits contain similar allegations of fraud and therefore must be treated as a single case study of enforcement. To collect the allegations of fraud by whistleblowers, I scraped the universe of press releases from the Department of Justice "Justice News" website that relate to Medicare and whistleblowing, from 1994 (the start of the archives) through 2014. From these press releases, I hand-coded lawsuits that contain similar allegations of misconduct in similar types of medical care into a single case study. For example, one case study is the misuse of the outlier payment system, which contains 11 press releases from different settlements with similar allegations. I omit case studies for which I do not have data, including enforcement that precedes the start of my data, or allegations related to falsification not visible in the Medicare claims. Appendix C describes this process in detail. Table 2 lists the 4 case studies with the largest total settlement amounts for which I have data, which comprise 29 press releases detailing $1.9 billion in total settlements. While presented as four case studies grouped by fraudulent conduct, these case studies reflect lawsuits against hundreds of providers, and represent about 7% of total healthcare whistleblower settlements. For each case study, I use court documents including whistleblower complaints and settlement agreements to gather details about the alleged conduct and guide the analysis of claims.

5.3 Case Details

5.3.1 Outlier Payment Falsification

The first case study concerns the misuse of outlier payments for inpatient hospitalization, for which over $900 million in settlements was recovered by the government between 2004 and 2010. Medicare pays providers of inpatient medical care a fixed reimbursement amount for the diagnosis related group (DRG) under which the patient is coded. By fixing reimbursement for each diagnosis, providers have incentives to keep costs down. However, this raises concerns that providers would be unwilling to treat high-cost patients. To mitigate this adverse selection effect, the Medicare system contains a provision for outlier payments, which are additional reimbursements for very-high-cost patients. Before 2004, to qualify for outlier payments, a patient must have exceeded a cost threshold, computed with a complicated

formula based on the provider's labor costs, capital costs, historic charges, and a geographic adjustment factor.[11] This formula provided an opportunity for misreporting: by manipulating their historic costs, hospitals were able to change their thresholds and collect more outlier payments.

On November 4, 2002, Tenet Healthcare, a large investor-owned hospital company, was sued under the False Claims Act for manipulating its cost reports in order to illicitly receive additional outlier payments.[12] This lawsuit was settled in June, 2006, with Tenet paying $788 million to resolve these allegations without admission of guilt. The DOJ press releases describe 10 other settlements for alleged manipulation of outlier payments. Appendix D.1 contains additional details about the related lawsuits.

Outlier payments constitute their own type of spending by the Medicare system, and are an accounting measure rather than a medical treatment *per se*. Therefore, for its controls, I consider the other broad types of payments made by Medicare that are of comparable scale, including durable medical equipment, home health care, hospice care, nursing care, and disproportionate share payments for hospitals that serve many low-income patients.

5.3.2 Medically Unnecessary Botox

The second case regards medically unnecessary usage of Botox. Despite popular branding as an "anti-wrinkle" procedure, Botox is FDA-approved for a number of important medical uses, including treatment of crossed eyes (strabismus) and neck spasms (cervical dystonia). Medicare covers medically necessary Botox injections for FDA-approved uses, but not for non-FDA-approved uses.

Between 2007 and 2009, Allergan, the sole manufacturer of Botox, was sued by a set of whistleblowers who alleged that Allergan had illegally promoted Botox for non-FDA-approved ("off-label") uses, including headaches. In order to ensure that Medicare would pay for the injections, Allergan allegedly instructed physicians to miscode the injections, using diagnosis codes for approved uses. Additional details about the outpatient coding of Botox and the whistleblower lawsuits are presented in Appendix D.2. On August 31, 2010, Allergan settled with the federal government for $600 million, of which $210 million was for federal civil liability (primarily Medicare overbilling), $375 million was a criminal fine, and $14.85 million was to recompense affected state Medicaid programs.

For the synthetic control design, Botox is compared to other outpatient procedures that saw similar pre-whistleblowing levels and trends in spending. Appendix D.2 contains additional details about these control units.

5.3.3 Unnecessary Inpatient Kyphoplasty

Kyphoplasty is a spine procedure to repair vertebral compression fractures that cause pain and deformity of the back, often observed among patients with osteoporosis. Kyphoplasty involves the percutaneous (through the skin) injection of bone cement into an inflatable balloon placed within the affected vertebra. Because the procedure is performed percutaneously, kyphoplasty can be safely conducted as an outpatient procedure. The kyphoplasty procedure was developed, patented, and marketed by the company Kyphon, which sold a spine surgery kit as well as other related medical devices (Kasper, 2010). Hospitals using the kyphoplasty procedure on Medicare patients would purchase the equipment from Kyphon and bill Medicare for the procedures that used these kits.

---

[11]Rawlings and Aaron (2005) provide a detailed analysis of this computation.

[12]One lawsuit against a different defendant, HealthSouth, which was filed in 1998 and settled in 2004, contained $89 million of settlement to resolve allegations of outlier payment manipulation. However, it appears the DOJ added the outlier allegations to an existing lawsuit against HealthSouth following the filing of the Tenet lawsuit after 2002. Therefore, I consider the Tenet lawsuit the first outlier lawsuit, and use its filing date as the treatment date. Around the time of filing, Tenet also received substantial negative press regarding its overuse of outlier payments; the timing of these reports, days before the filing of the lawsuit, may indicate that the whistleblowing case was leaked to investors.

In December, 2005, Kyphon was sued by FCA whistleblowers who alleged that Kyphon illegally promoted the procedure as an inpatient procedure as opposed to outpatient. By doing so, hospitals received greater reimbursement for the treatment, allowing Kyphon to charge more for its products. Hospitals kept patients for a short inpatients stay so they could receive the inpatient reimbursement level for a low amount of inpatient care. Inpatient stays under the relevant diagnosis-related groups (DRGs) were reimbursed in the $6,000 - $11,000 range, as opposed to outpatient kyphoplasty which was reimbursed between $500 and $2,000.

In May 2008, Kyphon settled these allegations with the Department of Justice for $75 Million, without admission of guilt. Between 2009 and 2015, the DOJ released another 9 press releases detailing settlements with 140 hospitals having performed unnecessary inpatient kyphoplasty. The sum of the settlements against Kyphon and the defendant hospitals was $214.2 Million. Appendix D.3 provides additional details for these lawsuits. I analyze inpatient spending on short stays of 7 or fewer nights under the inpatient DRGs promoted by Kyphon, and outpatient spending on all spine procedures. As controls for short-stay inpatient visits, I use inpatient spending for short stays of 7 or fewer nights under other DRGs that saw similar rises in spending. For controls on outpatient spine procedures, I use spending on other outpatient surgical procedures on the musculoskeletal system. Appendix D.3 describes the coding of inpatient and outpatient kyphoplasty and the control units used.

5.3.4 Unnecessary Inpatient Admission

The fourth case study concerns the unnecessary admission of Medicare beneficiaries for inpatient care at hospitals, instead of receiving observational or outpatient care. Many of these patients presented at the hospital's emergency department and should have been held under observational or outpatient status, which are reimbursed much less than inpatient care. The first successful lawsuit of this type was filed in October 2004, and in total, 7 settlements were reached regarding 135 hospitals for a total of $172.3 million in recovery between 2007 and 2014. The majority of the enforcement comes from the settlement with Community Health Systems, the nation's largest operator of acute care hospitals at the time, which settled for $98 million in 2014 for similar conduct in 119 of their hospitals. Appendix D.4 provides additional details about these lawsuits.

Unnecessary admissions were concentrated among certain hospitals, motivating an analysis of the defendants of these lawsuits rather than all inpatient care nationwide. The set of potential controls for the defendants are all other hospitals not litigated against for unnecessary inpatient procedures. To mitigate spillover effects into the control groups, I restrict the controls to hospitals in states that contained no defendants. These hospitals treat different patient pools than the defendants and are less likely to have doctors or administrators cross-employed with the defendant hospitals.

The outcome variable in this case study is total inpatient spending, and for the substitution effect, total outpatient spending. For each of the defendants, I construct a random sample of 100 control units, where each control group contains the same number of hospitals as the defendant. For example, two defendants were chains of 6 hospitals each; I create 100 control units of 6 randomly grouped control hospitals, drawn with replacement, from the set of controls. Similarly, to measure substitution by the defendant providers to increased outpatient expenditure, I use randomly grouped outpatient providers from the unaffected states. Appendix D.4 provides further details about this process.

5.4 Results

Figure 3 shows the main results of the synthetic control method. In 3 of the 4 case studies, all except Botox, whistleblowing caused a large decline in spending relative to the synthetic control unit, indicating strong specific deterrence effects of whistleblowing on fraudulent provider behavior. Each case study is analyzed at the month level, and each outcome variable is total payments from Medicare. In each case study, the pre-treatment fit of the treated unit on the controls is excellent, indicating that the synthetic control method successfully replicated the trends of the treated unit. Appendix Table A1 reports the Root Mean Square Prediction Error for each synthetic control analysis.

Table 3 summarizes the deterrence effects for these cases and provides totals, deterrence values, and deterrence-to-settlement ratios. These 4 whistleblower case studies produced deterrence effects totaling to $18.92 billion, compared to settlements of just $1.9 billion. The mean deterrence effect for these cases is 6.8 times the settlement value. There is substantial heterogeneity in the deterrence ratios, from a small negative deterrence effect in the Botox case to a particularly high positive deterrence ratio for the Outlier Payments case. Notably, the deterrence metric used here is computed using a discounted difference between the treated and control units for only five years after the filing of the case, giving 0 weight to deterrence effects beyond five years, and does not include general deterrence effects, i.e. spillovers to other types of fraud. Therefore, this is intentionally conservative in measuring the total deterrence effects of whistleblowing.

The largest of these effects is in the Outlier Payments case (top left in Figure 3): the 5-year discounted deterrence measurement for the outlier payments computed is $17.46 billion, which is roughly 19 times the total settlement value of the outlier whistleblowing lawsuits of $923 million. Visually, the synthetic control method estimates the rise of the outlier payments system at roughly a linear trend equal to its pre-period rise in the absence of whistleblower enforcement. The magnitude of the deterrence is driven by the scale of spending on outlier payments, which exceeded $500 million per month in its pre-whistleblowing peak, and then dropped off substantially following the lawsuits.

Notably, for the Botox case (top right in Figure 3), there is a small negative deterrence effect: Botox spending exceeds the synthetic control group post-lawsuit. The 5-year discounted deterrence effect is -$41.67 million, around 7% of the settlement value of $600 million. One potential reason for the negative deterrence effect is that Botox gained FDA approval for migraine coverage about 2 months after settling with the Department of Justice for illegally promoting botox for headaches (Singer, 2010). Because civil litigation and settlement negotiations can stretch out for indefinite periods of time, it is possible that Allergan timed the settlement to coincide with its expected FDA approval. This case exhibits that deterrence effects are not necessarily positive, and that the future value of misconduct is not necessarily large when compared to the past costs and settlement amount. In this circumstance, the $600 million settlement paid by Allergan to the US functioned as a penalty for promoting its product for a use that was not yet FDA approved. But given that FDA approval did ultimately arise, the future value of the damages and the specific deterrence effect are small.

For the Kyphoplasty case (bottom left) and Unnecessary Inpatient Admissions case (bottom right), Figure 3 shows that inpatient spending declined relative to the respective synthetic controls. The short-stay inpatient deterrence total for the Kyphoplasty case is $538.9 million. For the Unnecessary Inpatient Admissions case, I graph results for the Community Health Systems lawsuit, the largest defendant by far. Inpatient deterrence for the defendant Community Health Systems is $693.2 million, and the total inpatient deterrence for all defendants is $1.124 billion. Appendix Figure A2 shows the deterrence effect at other defendants. These decreases in inpatient spending must be weighed against expected increases in outpatient spending. Figure 4 plots the substitute outpatient spending for these cases. In the Kyphoplasty case, the increase in outpatient spending on all spine procedures totals to $257.8 million; when

compared with an inpatient spending decrease of $538.9 million, this results in a net deterrence effect of $281.1 million. For the Unnecessary Inpatient Admissions case, outpatient spending at the defendant CHS did not rise relative to the control providers. Appendix Figure A3 displays the similar synthetic control setup for each of the other defendants' outpatient spending, and shows heterogeneity, with some defendants' outpatient spending rising post-lawsuit and others' falling. The total deterrence from decreases in outpatient spending is $96.9 million.

Many of the particular control units that comprise the synthetic controls in these case studies were also due to fraudulent spending, which supports the validity of these controls to estimate the spending trajectory of the treated unit. Appendix Table A4 shows the synthetic control weights and time shifts for the control units for the Outlier case. Most of the synthetic control weight is placed on disproportionate share payments, with only a 1-month time shift. Disproportionate share payments operate very similarly to outlier payments in that they are additional payments for inpatient stays, and were also subject to later whistleblower lawsuits for overuse.[13] Appendix Table A5 shows the weights for the Botox case study among other outpatient CPT codes. Similarly, the top 2 controls by weight, retroperitoneal ultrasound and debridement, which were given 60.2% and 22.6% respectively by the synthetic control process, were also subject to enforcement for fraudulent overbilling at much later dates.[14] In the Kyphoplasty case, Appendix Table A6 shows that 43.1% of the synthetic control weight was placed on inpatient rehabilitation; as with the other case studies, rehabilitation was eventually the subject of anti-fraud enforcement due to overuse and improper billing.[15] The ultimate enforcement against the most heavily weighted control groups for fraud at later dates supports the validity of these control groups in estimating counterfactual spending without whistleblowing.

Overall, these results indicate that the specific deterrence benefits of whistleblowing cases often exceed the settlement values many times over, and greatly exceed the retrospective damages used to compute those settlement values. This indicates a large savings to the Medicare program as a result of these whistleblowing cases, exceeding both the recoveries to the government from the settlement as well as the whistleblower compensation.

5.5 Inference and Robustness

I conduct inference on the synthetic control deterrence estimates using permutation placebo testing, following Abadie et al. (2010). For each control, I construct a staggered synthetic control unit using all other control groups in the donor pool, and then construct the placebo deterrence value. I then compare the real estimated deterrence to the distribution of these placebo deterrence values. I conduct a 1-tailed test, which counts what fraction of placebos exceed the value of the treated unit's deterrence amounts, comparing positive deterrence values to other positive deterrence values and negative to negative.

Table 4 presents the results of the placebo test. These results indicate that the deterrence effects found are not due to chance. The deterrence total for the Outlier Payments case exceeds 100% of the placebo units. The small negative deterrence effect for Botox – that is, increased spending after whistleblowing – exceeds all but 3 of the 93 controls, indicating that this effect is statistically different from 0 despite the small magnitude. For the Kyphoplasty

---

[13]See, e.g., Department of Justice Office of Public Affairs (2009), for one such settlement.

[14]See Department of Justice Office of Public Affairs (2017) for a 2017 whistleblower settlement regarding unnecessary ultrasounds. Unnecessary retroperitoneal ultrasound usage specifically under CPT code 76775, the code given weight in the Botox synthetic control, was also the subject of 2018 provider self disclosure resulting in civil penalties (Office of the Inspector General of Health and Human Services, 2018). Debridement was also subject to overuse, motivating a 2007 report by the Office of the Inspector General of Health and Human Services (Office of the Inspector General of Health and Human Serivces, 2007), as well as a whistleblower lawsuit (United States *et al* ex. rel. Raalte *et al*, Docket 6:14-cv-00283 M.D. Fl.).

[15]See, e.g. Department of Justice Office of Public Affairs (2019), where the nation's largest inpatient rehabilitation provider was sued under the False Claims Act for unnecessarily admitting patients, as well as miscoding them, and settled for $48 million.

case, the reduction in inpatient spending exceeds 26 of the 30 placebos, and the corresponding increase in outpatient spending exceeds 14 of the 15 placebos. For the Unnecessary Inpatient Admissions case, there is strong evidence that the reduction in inpatient spending is not a chance finding; the 5 largest defendants (of 7) exceed between 93 and 99 of the 100 placebo units. However, substitution to outpatient spending shows mixed results, including statistically significant values in both the positive and negative direction. This mix of positive and negative effects indicates heterogeneity in how whistleblowing changed substitute outpatient spending at the defendant hospitals. Some hospitals may have reduced total volume after being sued, causing both inpatient and outpatient spending to decline, while others substituted from inpatient to outpatient spending. Taken together with the good pre-period fit shown in Figure 3, the placebo results indicate that the staggered synthetic control method measures large, statistically significant deterrence effects due to whistleblowing.

One potential concern about staggered synthetic controls is that controls that are shifted backward in time to match the treated unit could potentially be contaminated by the event in question. If the event contaminates the controls, the pre-period fit of the controls on the treated unit would be an invalid way of constructing an estimate of the counterfactual untreated series. To mitigate these concerns, I repeat the staggered synthetic control exercise with the additional restriction that controls are only shifted forward in time; that is, the trends of the controls used in fitting the pre-period occurred exclusively before the treatment date, and therefore cannot be contaminated by spillovers. Appendix Figure A4 presents the corresponding figure for this analysis. The figure is nearly indistinguishable from the original Figure 3, and the deterrence measurements are nearly identical, if only slightly greater than the original estimates. Therefore, spillovers onto the staggered controls is not a source of bias in this analysis.

The staggered synthetic control methodology differs from the traditional synthetic control methodology of Abadie et al. (2010) in that it assumes common time trends but not common calendar-time shocks. To mitigate concerns that this is a source of bias, I partial out time fixed effects (at the month level) for both the treated unit and its controls, and then re-estimate the staggered synthetic control model. Appendix Figure A5 presents the main effects of this analysis, and Appendix Figure A6 shows the effects on the substitute procedures. The total deterrence effects from this method are $27.6 billion, or around 1.5 times the main specification's estimated deterrence. The deterrence estimates of the Outlier Payments case are largely increased under this specification, while the deterrence estimates of the Kyphoplasty case are diminished. Qualitatively, Appendix Figure A5 shows good pre-period fits from the control units, and the same pattern of divergence between the synthetic control and the treated unit in the post-treatment period. Removing time fixed effects from the synthetic control estimator is similar in form to the demeaned estimator proposed by Ferman and Pinto (2021), who show that this can correct for bias due to unobserved confounding.

As a final robustness check, mitigating concerns about the appropriateness of the selected controls, I consider an estimation strategy without synthetic controls. I compute average spending in the last 12 months before the whistleblower files in each case, and compute deterrence as though spending were equal to this constant in the 5 years following whistleblowing. This intentionally understates a linear trend, which would have projected spending upward in the post-whistleblowing period for all cases. Appendix Figure A7 shows the result of this method on the main results. The total deterrence measurement under this method is $5.56 Billion, which includes the net increases in substitute procedures (figure not shown). Even under this conservative estimate, in which whistleblowing would not reduce expenditures but merely prevent them from continuing to rise, whistleblowing produces very large deterrence effects.

6   Estimating the Costs of Medicare Whistleblowing

Although False Claims Act litigation is conducted privately, whistleblowing incurs public and private costs, as discussed in the model above. While FCA litigation produces strong deterrence effects, this policy could be inefficient if it does so by incurring high legal costs associated with enforcement. Conversely, if FCA costs are low, it indicates that the FCA is a cost-effective way of combatting healthcare fraud. The goal of this section is to provide an estimate of the public costs of healthcare-related FCA cases and to contextualize these costs against the benefits of private enforcement and the costs of other enforcement mechanisms.

6.1   Public Costs

Overall, I estimate that public expenditure on the 445 healthcare-related False Claims Act cases filed in 2018 amounted to less than $108.5 million. Public expenditure on FCA occurs from a few different federal agencies: the Department of Justice, the Office of the Inspector General of Health and Human Services, and the federal courts. For each of these agencies, there are limitations to estimating costs expressly associated with the False Claims Act, because data are aggregated across multiple responsibilities of these agencies. To overcome this limitation, I estimate FCA-related spending by combining data on the legal process of FCA lawsuits, data on the number of FCA lawsuits, and the agencies' public budgets. Each of the amounts included are intentionally conservative in the direction of overestimation.

The Department of Justice, whose attorneys review and sometimes intervene in whistleblower cases, spent no more than $99.1 million on healthcare-related whistleblower FCA lawsuits in Fiscal Year 2018. In the Health Care Fraud and Abuse Control Program Annual Report for Fiscal Year 2018 from the Department of Justice and HHS to Congress, Department of Justice expenditures on overall healthcare fraud amounted to $135.3 million. However, these resources include $25.3 million for the criminal division and $10.9 million for the civil rights division, neither of which handle whistleblower lawsuits. This leaves $99.1 million for all other DOJ spending related to healthcare fraud, including all healthcare anti-fraud spending by the DOJ civil division, the US attorneys, and the FBI. While certainly much of these expenditures were for non-FCA cases, particularly FBI expenditures on criminal healthcare fraud, $99.1 million forms an upper bound of total healthcare whistleblower-induced spending by the DOJ in 2018.

False Claims Acts healthcare-related whistleblower lawsuits involve the Office of the Inspector General of Health and Human Services (OIG-HHS). The main relationship between OIG-HHS and the False Claims Act is through the Office of Counsel, a small internal department that provides general legal support to the Office and also oversees Corporate Integrity Agreements for companies settling False Claims Act lawsuits. The total expenditure of the Office of the General Counsel was $7.1 million 2018 (The Department of Health and Human Services And The Department of Justice, 2019). Although some of these expenditures support non-FCA responsibilities of the Office of Counsel, I use $7.1 million as an upper bound on OIG-HHS spending on these cases.

I estimate that the Federal Courts, which try the whistleblower lawsuits, spent $2.3 million in 2018 on FCA cases. While there are no data available from the Courts on case-specific spending, the average cost of lawsuits is an upper bound for the marginal cost of lawsuits. In 2018, the federal courts spent a total of $7.7 billion on all operations, handling 1.48 million federal lawsuits across a variety of topics, including criminal prosecutions, bankruptcy, and the court of international trade (The Administrative Office of the US Courts, 2019). The Department of Justice reported 446 healthcare-related lawsuits in 2018 (US Department of Justice, 2018) or about 0.03% of the total federal cases. Therefore, the aggregate cost associated with these cases can be estimated at $0.0003 \times \$7.7\text{Bil} = \$2.3\text{Mil}$. Given that the federal court system handles nearly a million and a half lawsuits, the marginal costs associated with these additional

445 cases are potentially even smaller, as they benefit from use of an established court system largely devoted to other areas of law.

Overall, the total expenditures of the federal government on False Claims Act whistleblower healthcare lawsuits were no more than $108.5 million in FY2018. This expenditure is small compared to the deterrence effects of just the few largest False Claims Act cases, and also to government expenditure on top-down antifraud enforcement. In Fiscal Year 2018, total federal healthcare fraud and abuse resources across agencies amounted to $2.04 billion, of which the vast majority was unrelated to civil False Claims Act enforcement (The Department of Health and Human Services And The Department of Justice, 2019). The largest budget item by far is the Medicare Integrity Program (MIP) at $809 million in FY 2018, focusing on top-down fraud identification including audits and medical reviews. In comparison to the large anti-fraud expenditure, whistleblowing is a particularly low-cost way to combat and deter healthcare fraud, and produces extensive public benefit.

Whistleblower payouts also appear as an accounting liability to the government when considering the costs of the False Claims Act. However, these payments do not affect social welfare, as discussed in Section 3, because they are transfers from the defendant to the whistleblower. Total whistleblower payouts for healthcare-related cases in my data total to $4.29 billion, which is a relatively small figure compared to the tens of billions of dollars in recovery and deterrence that these cases have produced. Statistics focused on whistleblower payouts, which are regularly included as a "cost" in government accounting, misconstrue the nature of these payments which are not in fact a public cost at all.

6.2 Private Costs

Another expense contributing to the efficiency of the False Claims Act are the private costs of whistleblower lawsuits, captured in the model as C_p. Private costs for plaintiffs and defendants are difficult to measure accurately, as there are no public data sources that compile this information. However, ballpark figures can be gleaned from other models of civil litigation costs that use surveys of attorneys' hours and expenses to estimate costs. A 2013 study estimated that the median professional malpractice lawsuit cost $122,140, the highest of all surveyed categories, while the median automobile tort lawsuit cost only $43,238, the lowest category (Hannaford-Ago, 2013). Even if False Claims Act cases cost *double* the reported price of the average malpractice lawsuit for both plaintiff and defendant, the 446 healthcare-related whistleblower lawsuits filed in 2018 would cost only an estimated $109 million. Therefore, both the public and private costs of whistleblowing lawsuits are dominated by the benefits of recovered funds and deterred overspending.

7 Measuring Changes in Patient Care

In addition to the fiscal effects described in the previous section, whistleblowing under the FCA creates incentives for providers to change the way they treat patients. These changes could be either beneficial or harmful for patient health: if whistleblowing curbs behavior that was profitable to providers at the expense of patient health, then we expect whistleblowing to benefit patients. However, whistleblowing could also induce defensive behavior among physicians, influencing their care decisions away from what is beneficial to patients and instead to what would be justifiable if they were sued. This echoes similar concerns about defensive medicine from the medical malpractice literature, notably Kessler and McClellan (1996), as well as evidence that increased disclosure from hospital "report cards" led to increased provider selection and patient sorting (Dranove et al., 2003).

I conduct a measurement exercise to understand whether patient care was improved or harmed by whistleblowing. As a case study, I examine provider care decisions following changes in kyphoplasty due to the whistleblower case described in Section 5.3.3. Kyphoplasty is an ideal case study for the discussion of provider care decisions for a few reasons. First, there was a large reduction in inpatient usage of this procedure and a substitution to outpatient procedures, indicating a change in actual care decisions by providers. This is in contrast to the Outlier Payments case study, which seems to be a change in billing procedures, or to the Botox case, where there was little effect on usage. Second, kyphoplasty is a single procedure with previously studied health effects (discussed below), allowing for a targeted analysis of the effects of whistleblowing on patient care. This is in contrast to the unnecessary inpatient admissions case, which related to a broad set of medical procedures. As such, kyphoplasty is the largest-settlement-value case study for which I can conduct an analysis of provider care decisions.

As discussed in Subsection 5.3.3, Kyphoplasty is a spine procedure to repair compressed vertebrae in patients with osteoporosis. Kyphoplasty can be very beneficial for patient health: a recent meta-analysis of vertebral compression fractures (VCFs) shows that patients with VCFs have 2.5 times the mortality rate of patients without them, and that kyphoplasty is successful at reducing mortality rates compared with non-operative care (Kurra et al., 2018). One potential mechanism is that vertebral compression fractures can compromise pulmonary function, leading to greater rates of pneumonia, and surgical intervention corrects this issue Chen et al. (2013). Estimates in the meta-analysis range from 35% to 70% mortality reduction over a 3 to 5 year period after receiving kyphoplasty, indicating a potentially valuable mortality reduction from this procedure. Still, Kyphoplasty remains a controversial procedure among the medical community due to the heterogeneity of its effectiveness found by different studies, and its potential overuse among inappropriate patients. The overall effectiveness of Kyphoplasty is beyond the scope of this paper; instead, I focus on the selection of patients into receiving treatment, based on whether those individual patients are likely to have benefited from the procedure.

Whistleblowers alleged that Kyphon, the maker of the kyphoplasty surgical kit, encouraged short inpatient stays for the surgery. Kyphoplasty can be safely performed outpatient, but inpatient stays allowed hospitals to receive greater reimbursement and for Kyphon to charge more for the surgery kit. Short-stay inpatient treatment for kyphoplasty was drastically reduced following whistleblowing, with unknown effects on patient health. Some of the reduction in inpatient care involved a substitution to outpatient kyphoplasty, or to vertebroplasty, a close substitute procedure.

To understand the impact of these changes, I model patient death as a function of receiving kyphoplasty within a heterogeneous treatment effects framework. The goal of this exercise is to measure how provider care decisions changed due to whistleblowing, based on whether the patient is expected to benefit from the procedure. First, I construct two non-overlapping cohorts of patients from before and after the kyphoplasty lawsuit. For the purposes of the treatment effects model, I define "treatment" as receiving a 1-night inpatient stay in 2005 under the DRGs allegedly promoted by Kyphon.[16] My 2005 cohort includes every 70-75 year old treated for the first time in 2005 and the full population of never-before-treated 70-75 year old control Medicare patients. My 2011 sample similarly contains every 70-75 year old treated for the first time in 2011 and the full population of never-before-treated 70-75 year old control patients. These cohorts are non-overlapping, and therefore all 2011 cohort members are not potential controls for the 2005 cohort. The 2005 cohort consists of 8.2 million patients, and the 2011 cohort consists of 9.3 million patients.

For each patient, I collect extensive data, including inpatient or outpatient kyphoplasty/vertebroplasty treatment,

---

[16]In contrast to this setup, the spending amounts used to compute deterrence in Section 5 used all short stays (7 nights or fewer) under these DRGs and all outpatient spine spending. This potentially lumped in some non-kyphoplasty treatment, which was unaffected by whistleblowing and was differenced out when computing deterrence values. To compute health effects, I focus on 0 or 1-night stays under these DRGs, which almost completely vanished post-whistleblowing, as I have greater confidence that kyphoplasty was performed during these inpatient visits. Correspondingly, outpatient treatment is restricted to the outpatient codes specifically for kyphoplasty and vertebroplasty.

all inpatient claims, chronic conditions, and demographic data. Patient covariates include age, state, sex, race, original and current reasons for Medicare qualifications (i.e. age or disability). Inpatient claim data were taken from the 100% MedPar files; patients receiving short-stay kyphoplasty treatment at any time before the cohort year are excluded. As covariates, I also collect inpatient stay data for 6 years before the cohort year, i.e. 1999–2004 for the 2005 cohort and 2005–2010 for the 2011 cohort, and include an indicator for any inpatient stay, the number of stays, the patient's total inpatient stay length, a count of the number of stays under each DRG, and the total Medicare payment amount for that patient's inpatient treatment. Furthermore, I include chronic condition indicators for each patient in the cohort year, which detail a patient's chronic conditions such as Alzheimer's, hip fractures, or osteoporosis. Finally, for each patient I collect death dates, and produce an indicator of whether the patient died within 5 calendar years, i.e. 2005–2010 for the 2005 cohort and 2011–2016 for the 2011 cohort. The length of the claims history and death data used and the restriction to the 70-74 year age bin are due to the availability of data, which span from 1999 – 2016, to produce a maximally non-overlapping sample.

I estimate a regression where treatment is interacted with patient characteristics to examine how treatment benefits vary with observable characteristics. I consider mortality as the outcome, using short-stay inpatient treatment among the 2005 cohort, with the following logistic regression:

$$Death_i = \alpha + \beta T_i + \gamma' C_i + \delta' T_i C_i + \eta' M_i + \theta' T_i M_i + \varepsilon_i \tag{5}$$

The outcome variable $Death_i$ is an indicator if patient i died within 5 years. This limited dependent variable motivates a logit framework for the regression. T_i is the treatment indicator, C_i is the matrix of patient covariates and M_i is the matrix of patient claims history and chronic conditions. This specification models death in terms of treatment interacted fully with these controls. As such, the fitted model captures the effect of each covariate and each aspect of medical claims history on death, with or without short-stay inpatient kyphoplasty treatment. Appendix Table A2 presents select coefficients of this regression.

Using this model fitted to the 2005 sample, I can then predict the effects of kyphoplasty among both 2005 and 2011 patients. I construct: $\widehat{Y_{i1}} = P(Death|M_i, C_i, T_i = 1)$ and $\widehat{Y_{i0}} = P(Death|M_i, C_i, T_i = 0)$ for each patient, using the regression coefficients fit to the 2005 sample. I then produce a predicted treatment effect for each patient:

$$\widehat{TE_i} = \widehat{Y_{i1}} - \widehat{Y_{i0}} = P(Death|M_i, C_i, T_i = 1) - P(Death|M_i, C_i, T_i = 0) \tag{6}$$

This model makes the standard conditional independence assumption: that conditional on a rich set of controls, here inpatient claims history, chronic conditions, and patient covariates, that potential outcomes Y_{1i} and Y_{0i} under treatment or non-treatment are independent of actual treatment status. That is, by controlling for the factors that influence probability of treatment, one can construct both potential outcomes for each patient, despite only ever observing Y_1 or Y_0 for any given patient. Appendix Figure A10 shows the histogram of predicted treatment effects for patients in 2005 and 2011. These histograms exhibit a similar shape between the cohorts, which means the comparison between these cohorts is between like populations.

Figure 5 plots the probability of short-stay inpatient treatment by predicted treatment effect in each cohort.[17] The predicted treatment effect on the horizontal axis is the change in mortality from receiving treatment versus not receiving treatment. Units to the left of 0 are predicted to have reduced mortality if treated, while units to the right of 0

---

[17]To satisfy Medicare data cell-size suppression rules against reporting any result with $n < 11$, the bins in Figure 5 have been top coded at 0.4 and -0.2 to achieve a minimum number of treated units within the extreme bins.

have increased mortality if treated. The ideal targeting of treatment to patients who benefit would place all of the mass to the left of 0. In both 2005 and 2011, patients who stood to benefit from the procedure were about twice as likely to receive treatment. The comparison between 2005 and 2011 shows that there was a reduction in inpatient probability treatment across the spectrum of treatment effects, both for those harmed and helped by the treatment.

Total inpatient treatment volume for kyphoplasty was counteracted by substitution to outpatient treatments. Figure 6 plots the probability of receiving an outpatient kyphoplasty or vertebroplasty within each group. Because these procedures are similar whether performed inpatient or outpatient, differing mostly in terms of billing, the predicted inpatient treatment effect on the horizontal axis is a reasonable way of understanding the effect of having had the procedure in either location. In both cohorts, the probability of receiving treatment is higher for those helped by the treatment, to the left of the distribution. Between 2005 and 2011, outpatient treatment probability grew for all types of patients, but substantially more for patients for whom kyphoplasty is expected to reduce mortality.[18]

To examine the net effect of the substitution from inpatient to outpatient procedures, I examine the probability of either inpatient or outpatient treatment by heterogeneous treatment effect. Figure 7 breaks the population into two categories: those with reduced mortality if they receive the procedure (negative value treatment effect) and those with increased mortality (positive valued treatment effect). The results show an overall decrease in treatment probability to those harmed by the procedure between 2005 and 2011, and an increase in treatment probability to those who benefit from the treatment across the same time period. Those helped by the procedure saw an increase from 0.144% to 0.155% probability of treatment, a 7.6% increase. The small raw percentages reflect the fact that the analysis sample is the entire never-before-Medicare treated 70-75 year old population in these years, and that kyphoplasty is relatively rare. Correspondingly, patients who were expected to by harmed by the procedure saw a decrease in probability of treatment from 0.0547% to 0.0506%, a decrease of 7.4%. Appendix Figure A12 breaks this same analysis into finer groups, and plots the probability of receiving either inpatient kyphoplasty or outpatient kyphoplasty or vertebroplasty by heterogeneous treatment effects. There was, in general, an increase in the probability of receiving treatment between 2005 and 2011 for those who benefit from the procedure, and a decrease in treatment for those harmed by the procedure. As a caveat, this analysis uses a coarse measure of patient health, mortality, which does not capture changes to patient quality of life from receiving treatment. Furthermore, an analysis of this type does not easily lend itself to inference, and no inference is conducted here; however, the positive sign of the results allows us to at least rule out a strong negative effect.

I estimate that the number of lives saved by changes in who was treated, in just the 2011 cohort of 70-74 year olds, was 45 fewer deaths. As a baseline, there were 5,243 patients who received inpatient kyphoplasty or outpatient substitutes in the 2011 cohort. I compute the number of lives saved by integrating the change in treatment probability times the predicted treatment effect across the population:

$$\Delta Deaths = \sum_k \Delta P(Treated)_k * \widehat{TE}_k * N_k \tag{7}$$

Here, k indexes the treatment effect bins, $\Delta P(Treated)$ is the change in probability of treatment from 2005 to 2011, and N_k is the number of people in each bin in 2011. I use bins of size 0.001 to minimize error due to binning and most closely approximate an integral.

These results are consistent with the better targeting of kyphoplasty following the whistleblowing cases that reduced the volume of care. Patients who benefit from kyphoplasty were more likely to receive the procedure following

---

[18]Similar to Figure 5 , bins in Figure 6 have been top coded to achieve a minimum number of treated units within the extreme bins, in compliance with Medicare cell-size suppression rules.

the lawsuit, and those harmed were less likely to receive the procedure. One potential explanation is the change in incentives for providers, who before the lawsuit were more profit-motivated in their treatment, picking low-cost patients to receive procedures that could be heavily reimbursed, with less focus on the patient's expected health outcomes. Under this explanation, whistleblowing refocused provider attention on expected patient health outcomes, creating better targeting toward individuals who benefit the most. These effects, while small, assuage concerns that changes in litigation risk might be detrimental for patient health. This finding reflects Kessler and McClellan (1996) and Kessler and McClellan (2002), which find that provider malpractice litigation did not have negative health consequences for patients.

Overall, in the case of kyphoplasty, whistleblowing seems to have had positive effects on patients by inducing better targeting of the procedure to those who benefit from it. This evidence indicates that kyphoplasty was overused in 2005, before the lawsuit, as evidenced by treatments performed on those expected to be harmed by the procedure. Whistleblowing enforcement was successful at reducing treatment to those individuals as well as increasing treatment to individuals likely to be helped by the procedure. In this case, the positive effects of whistleblowing went beyond financial benefits to the Medicare program, and indeed had small positive effects for patient care.

8 Conclusion

Private enforcement is a potentially valuable way to improve the federal provision of services and eliminate waste, fraud, and abuse. The United States relies upon a private enforcement regime to conduct a major form of federal anti-fraud enforcement, whereby whistleblowers conduct lawsuits on behalf of the federal government in exchange for a share of the funds they recover. Many of these lawsuits have been related to Medicare and the federal health insurance programs, which are particularly susceptible to fraud. Privatization comes with trade-offs that are not fully internalized by the enforcers: whistleblowing has the potential for large deterrence effects, but may impose costs on the government, private firms, and people receiving public services. The efficiency of this system depends on the relative magnitudes of these values.

This paper models the trade-offs of whistleblowing and quantifies its effects using data from Medicare and the Department of Justice. I undertake a set of case studies of large whistleblower lawsuits and measure specific deterrence effects, the change in the type of spending a whistleblower indicated was fraudulent. I analyze four case studies for which whistleblowers recovered a total of $1.9 billion in federal funds. I estimate that these lawsuits generated $18.92 billion in specific deterrence effects. In contrast, public costs for all lawsuits filed in 2018 amounted to less than $108.5 million, and total whistleblower payouts for all cases since 1986 have totaled to $4.29 billion. Just the few large whistleblowing cases I analyze have more than paid for the public costs of the entire whistleblowing program over its lifespan, indicating a very high return on investment to the FCA.

Changes in medical care induced by whistleblowers can have effects on patient health. I model the health effects of kyphoplasty, a spine procedure for patients with osteoporosis that was affected by whistleblower lawsuits against more than a hundred hospitals. I find small beneficial changes to patient care, with better targeting to patients who are expected to benefit from this procedure. This case study motivates further analysis of the effects of whistleblowing on patient care. In addition, whistleblowing generates changes to the care of patients that are potentially unrelated to the quality of the provider or the procedure, and this may provide experimental variation that other researchers find useful in the analysis of medical outcomes.

Whistleblowing has other potential costs and benefits not quantified in this paper. The risk of litigation may cause

providers to forgo misreporting in the first place, particularly when whistleblowers are empowered to directly sue for their own profit. These general deterrence effects are hard to measure without knowing the types of potential fraud that could have been committed. The deterrence effects presented here are lower bounds of the total deterrence effects due to these spillovers, and therefore the total deterrence may be much greater. Conversely, increased compliance requirements impose costs on providers that are not measured here, and I only able to broadly estimate the private costs of whistleblower lawsuits.

In this paper I estimate the fiscal benefits of privatized enforcement as compared to the absence of such enforcement. However, a different counterfactual would be better public enforcement. Paying whistleblowers 15-30% of recovered funds is expensive if the government could produce similar recoveries without whistleblowing. Given the vast amount of data collected by the Medicare program, some of the effects of whistleblowing could likely be accomplished through machine learning, pattern detection, and automated audits. The fact that these programs are not yet in place may point to the limited enforcement capacity of the federal bureaucratic institutions.

The results of this analysis suggest that privatization is a highly effective way to combat fraud. Whistleblowing and private enforcement have strong deterrence effects and relatively low costs, overcoming the limited incentives for government-conducted anti-fraud enforcement. A major benefit of the False Claims Act is not just the information provided by the whistleblower, but also the profit motive it provides for whistleblowers to root out fraud.

References

31 US Code Section 3731. False claims procedure, 1986.

AAPC Coder. Surgical procedures on the musculoskeletal system CPT code range 20100-29999. `https://coder.aapc.com/cpt-codes-range/230`, 2019. Accessed: September 13, 2019.

Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132, March 2003.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105 (490):493–505, 2010.

Administrative Office of the US Courts. The judiciary fiscal year 2018 congressional budget summary, 2017. Accessed: August 20, 2019.

Kenneth J. Arrow. Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5): 941–973, 1963.

Susan Athey and Guido W Imbens. Design-based analysis in difference-in-differences settings with staggered adoption. Working Paper 24963, National Bureau of Economic Research, August 2018.

Lucian Arye Bebchuk and Howard F. Chang. An analysis of fee shifting based on the margin of victory: On frivolous suits, meritorious suits, and the role of rule 11. *The Journal of Legal Studies*, 25(2):371–403, 1996.

Lucian Arye Bebchuk and Alon Kelement. Negative expected value suits. *The New Palgrave Dictionary of Economics and the Law*, pages 551–554, 1998.

David Becker, Daniel Kessler, and Mark McClellan. Detecting medicare abuse. *Journal of Health Economics*, 24(1): 189–210, 2005.

Gary S. Becker and George J. Stigler. Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, 3(1):1–18, 1974.

Eli Ben-Michael, Avi Feller, and Jesse Rothstein. Synthetic controls and weighted event studies with staggered adoption. *arXiv*, 2019.

Eli Ben-Michael, Avi Feller, and Jesse Rothstein. The augmented synthetic control method. *arXiv*, (1811.04170), 2020.

Center for Medicare and Medicaid Services. Medicare 100% research identifiable files, 1999-2016.

Center for Medicare and Medicaid Services. List of diagnosis related groups (DRGS), FY 2005, 2005. Accessed: August 20, 2019.

Chamber of Commerce of The United States of America. Amicus Brief, State Farm Fire and Casualty Company v United States ex rel Cori Rigsby; Kerri Rigsby, 15-513 (Supreme Court 2015). 2015.

Andrew T Chen, David B Cohen, and Richard L Skolasky. Impact of nonoperative treatment, vertebroplasty, and kyphoplasty on survival and morbidity after vertebral compression fracture in the medicare population. *JBJS*, 95 (19):1729–1736, 2013.

Congressional Budget Office. Health care. https://www.cbo.gov/topics/health-care, 2019. Accessed: August 19, 2019.

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

Department of Justice, Civil Division. Freedom of information act request 145-foi-15832, 2018. Fulfilled April 9, 2018.

Department of Justice Office of Public Affairs. San Mateo county, California, to pay us $6.8 million to resolve false claims allegations, March 2009.

Department of Justice Office of Public Affairs. "Sarasota physician agrees to pay $1.95 million to resolve false claims act allegations regarding unnecessary ultrasounds", December 2017.

Department of Justice Office of Public Affairs. Encompass health agrees to pay $48 million to resolve false claims act allegations relating to its inpatient rehabilitation facilities, June 2019. URL https://www.justice.gov/opa/pr/encompass-health-agrees-pay-48-million-resolve-false-claims-act-allegations-relating-its.

Simeon Djankov, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. The law and economics of self-dealing. *Journal of Financial Economics*, 88(3):430 – 465, 2008. ISSN 0304-405X. Darden - JFE Conference Volume: Capital Raising in Emerging Economies.

David Dranove, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. Is more information better? the effects of "report cards" on health care providers. *Journal of political Economy*, 111(3):555–588, 2003.

Uwe Dulleck and Rudolf Kerschbamer. On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 44(1):5–42, March 2006.

David Freeman Engstrom. Harnessing the private attorney general: Evidence from qui tam litigation. *Columbia Law Review*, 112(6):1244–1325, 2012.

David Freeman Engstrom. Public regulation of private enforcement: Empirical analysis of DOJ oversight of qui tam litigation under the false claims act. *Northwestern University Law Review*, 107(4):1689–1756, 2013.

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

Bruno Ferman. On the properties of the synthetic control estimator with many periods and many controls. *arXiv*, (1906.06665), 2020.

Bruno Ferman and Cristine Pinto. Synthetic controls with imperfect pre-treatment fit. *arXiv*, (1911.08521), 2021.

Paula Hannaford-Ago. Measuring the cost of civil litigation: Findings from a survey of trial lawyers. *Voir Dire*, 2013.

Jonas Heese. The role of overbilling in hospitals' earnings management decisions. *European Accounting Review*, 27, 2018.

Nathaniel Hendren and Ben Sprung-Keyser. A unified welfare analysis of government policies. *Quarterly Journal of Economics*, 135(3):1209–1318, 2020.

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

David H Howard and Nihar Desai. False claims act investigations of unnecessary percutaneous coronary interventions. *JAMA Internal Medicine*, 180:1534–1536, September 2020.

Howell E. Jackson and Mark J. Roe. Public and private enforcement of securities laws: Resource-based evidence. *Journal of Financial Economics*, 93(2):207 – 238, 2009. ISSN 0304405X.

D. M. Kasper. Kyphoplasty. *Seminars in Interventional Radiology*, 27(2):172–184, Jun 2010.

Daniel Kessler and Mark McClellan. Do doctors practice defensive medicine? *Quarterly Journal of Economics*, 111 (2):353 – 390, 1996.

Daniel P. Kessler and Mark B. McClellan. How liability law affects medical productivity. *Journal of Health Economics*, 21(6):931 – 955, 2002. ISSN 0167-6296.

S. Kurra, U. Metkar, I. H. Lieberman, and W. F. Lavelle. The Effect of Kyphoplasty on Mortality in Symptomatic Vertebral Compression Fractures: A Review. *International Journal of Spine Surgery*, 12(5):543–548, Oct 2018.

David Kwok. Evidence from the false claims act: Does private enforcement attract excessive litigation? *Public Contract Law Journal*, 42(2):225–249, 2013. ISSN 00333441, 21628181.

William Landes and Richard Posner. The private enforcement of law. *Journal of Legal Studies*, 4(1), 1975.

Latham and Watkins White Collar Defense and Investigations Practice. Seventh circuit deepens circuit split over fca dismissal authority, August 2020.

Office of the Inspector General of Health and Human Serivces. Medicare payments for surgical debridement services in 2004. Technical Report OEI-02-05-00390, 2007.

Office of the Inspector General of Health and Human Services. "provider self-disclosure settlements", 2018. URL https://oig.hhs.gov/fraud/enforcement/cmp/psds.asp.

Thomas F III O'Neil, Adam H Charnes, and Shannon Thee Hanson. The buck stops here: Preemption of third-party claims by the false claims act. *The Journal of Contemporary Health Law and Policy*, 12(1):41–65, 1995.

A. Mitchell Polinsky. Private versus public enforcement of fines. *The Journal of Legal Studies*, 9(1):105–127, 1980.

R. Brent Rawlings and Hugh E. Aaron. The effect of hospital charges on outlier payments under Medicare's inpatient prospective payment system: Prudent financial management or illegal conduct? *Annals of Health Law*, 14(2): 267–328, 2005.

Steven Shavell. The social versus the private incentive to bring suit in a costly legal system. *The Journal of Legal Studies*, 11(2):333–339, 1982. ISSN 00472530, 15375366.

Steven Shavell. Specific versus General Enforcement of Law. *Journal of Political Economy*, 10 1991. ISSN 1088-1108.

Elaine Silverman and Jonathan Robert. Skinner. Medicare upcoding and hospital ownership. *Journal of health economics*, 23 2:369–89, 2004.

Natasha Singer. Botox shots approved for migraine. *The New York Times*, Oct 2010.

Kathryn E. Spier. Chapter 4: Litigation. In A.M. Polinsky and S. Shavell, editors, *Handbook of Law and Economics*, volume 1 of *Handbook of Law and Economics*, pages 259 – 342. Elsevier, 2007.

The Administrative Office of the US Courts. The judiciary fiscal year 2020 congressional budget summary. February 2019.

The Department of Health and Human Services And The Department of Justice. Health care fraud and abuse control program annual report for fiscal year 2018. May 2019.

United States Department of Justice. Justice department celebrates 25th anniversary of false claims act amendments of 1986. https://www.justice.gov/opa/pr/justice-department-celebrates-25th-anniversary-false-claims-act-amendments-1986, January 2012. Accessed: August 19, 2019.

United States District Court, D.C. *US ex rel. Rockefeller v. Westinghouse Electric Co, 274 F.Supp.2d 10*. 2003.

United States Government Accountability Office. Report to congressional committees: High risk series. substantial efforts needed to achieve greater progress on high-risk areas. GAO-19-157SP, March 2019.

US Department of Justice. Civil division, fraud statistics. https://www.justice.gov/civil/page/file/1080696/download, 2018. Accessed: August 28, 2019.

Figures and Tables

Figure 1: Specific Deterrence
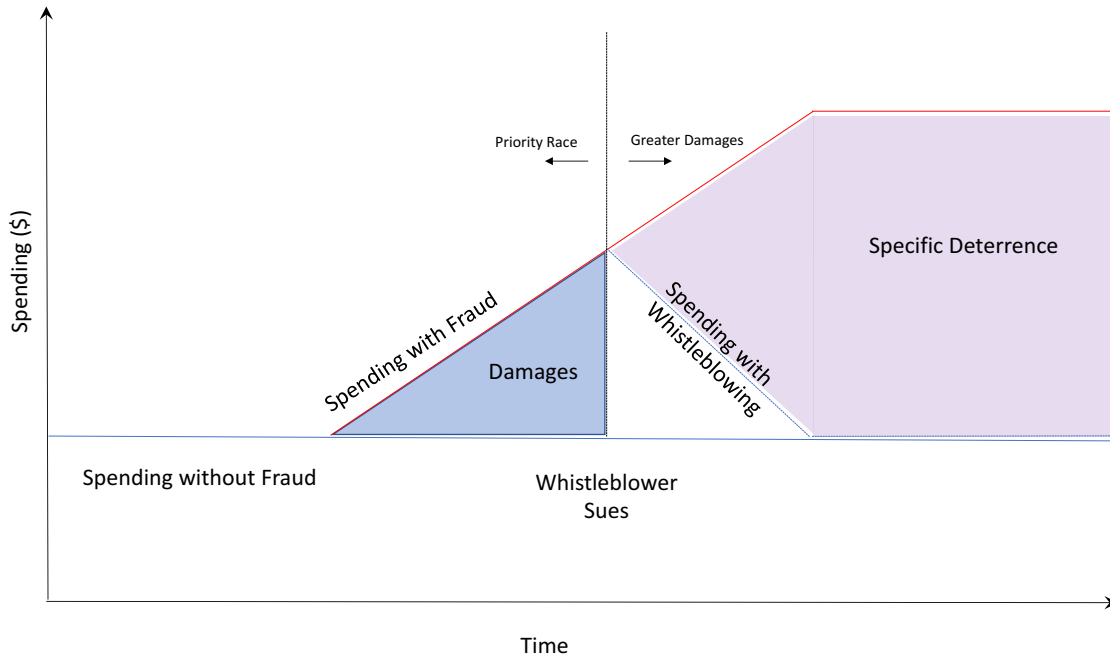
Notes: This figure describes the theoretical effects of a successful whistleblowing case on federal spending. When fraud is committed, the government has damages that are the difference between spending with fraud and the counterfactual spending without fraud. After the whistleblower sues, spending decreases back to its pre-fraud levels. Time trends are presented as linear trends for simplicity. Without loss of generality, fraudulent spending rises and eventually asymptotes, as it cannot grow infinitely even in the absence of enforcement. The specific deterrence effect is the difference between how much would have been spent without whistleblowing and how much is spent after whistleblowing occurs. Because whistleblowers are paid proportionally to the damages, they have incentives to blow the whistle later and allow the damages to accumulate; however, because the first whistleblower to come forward receives greater compensation, they have countervailing incentives to file as soon as possible.

Figure 2: Example of Staggered Synthetic Controls
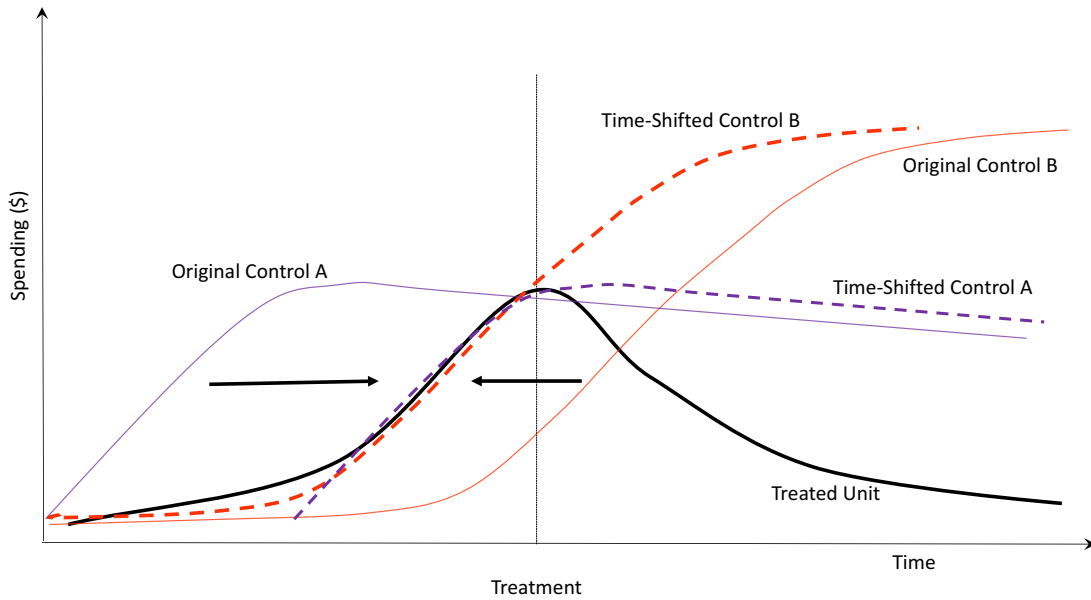
Notes: This figure exemplifies the fitting process for staggered synthetic controls. Spending on the treated unit is a solid black line that increases pre-treatment and decreases post-treatment. Control A exhibits a similar rise to the pre-period, but at an earlier time, and is shifted forward. Control B exhibits a comparable rise at a later period, and is shifted backward. The shifts are picked to best approximate the pre-treatment period in both shape and levels. These fits are agnostic to how the controls develop in the post-treatment period; Control A falls while Control B continues to rise. Following these fits, a synthetic control unit can be constructed from Time-Shifted Control A and Time-Shifted Control B.

Figure 3: Effects of Whistleblowing on Medicare Expenditure
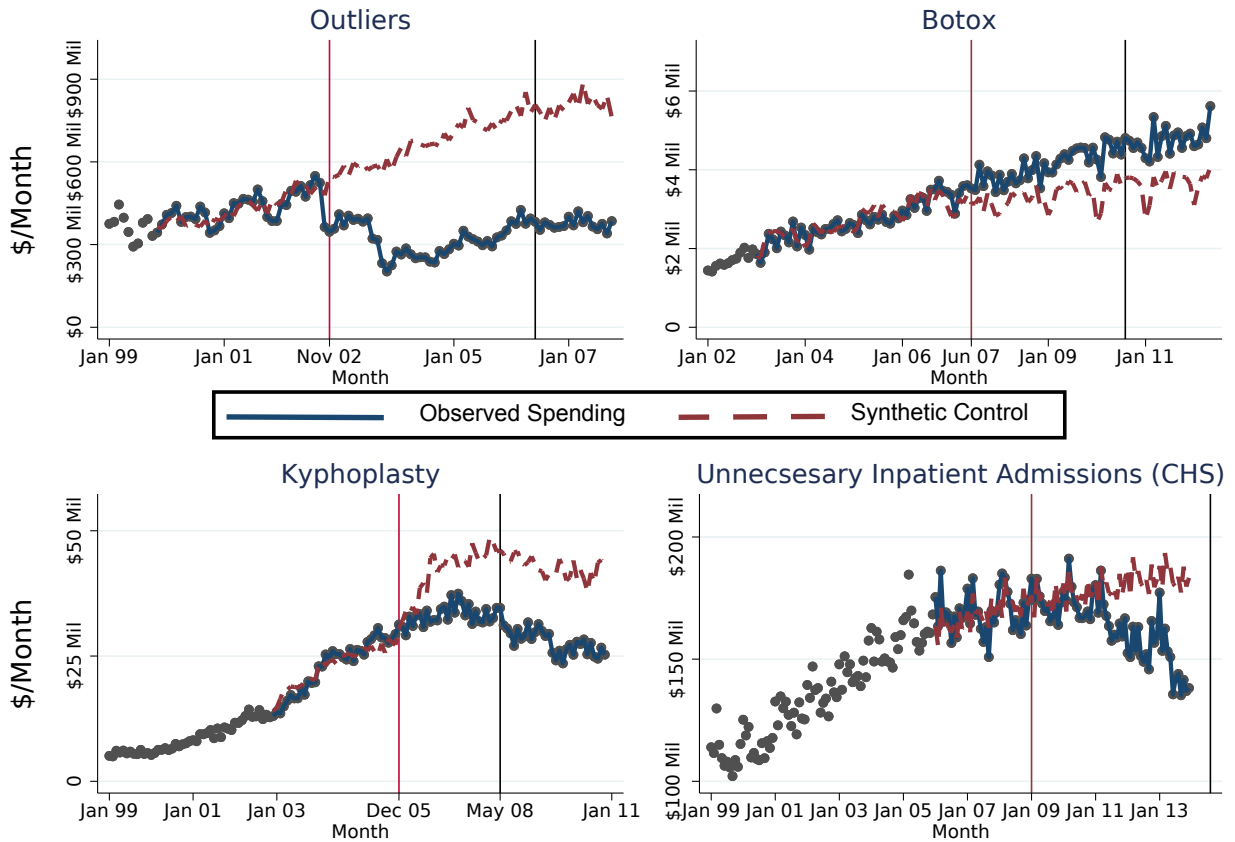
Notes: This figure plots the main effects of the 4 case studies: Outlier Payments (top left), Botox (top right), Kyphoplasty (bottom left), and Unnecessary Inpatient Admissions (bottom right). For each case, the spending affected by whistleblowing is plotted in blue, while the synthetic control series is plotted in red. The grey dots represent the spending on the treated unit in the period before it overlaps with the synthetic control group. The first vertical line of each case represents the filing of the first related whistleblower lawsuit, which is used as the treatment date, and the second vertical line reflects the first settlement. Post-treatment effects are analyzed for 5 years after the treatment date. For the Unnecessary Inpatient Admissions Case (bottom right), multiple defendant hospitals were analyzed, and the series included here reflects Community Health Systems, the largest defendant hospital chain. Appendix Figure A2 plots the same figure for the other defendants in that case.

Figure 4: Synthetic Controls for Substitute Outpatient Spending
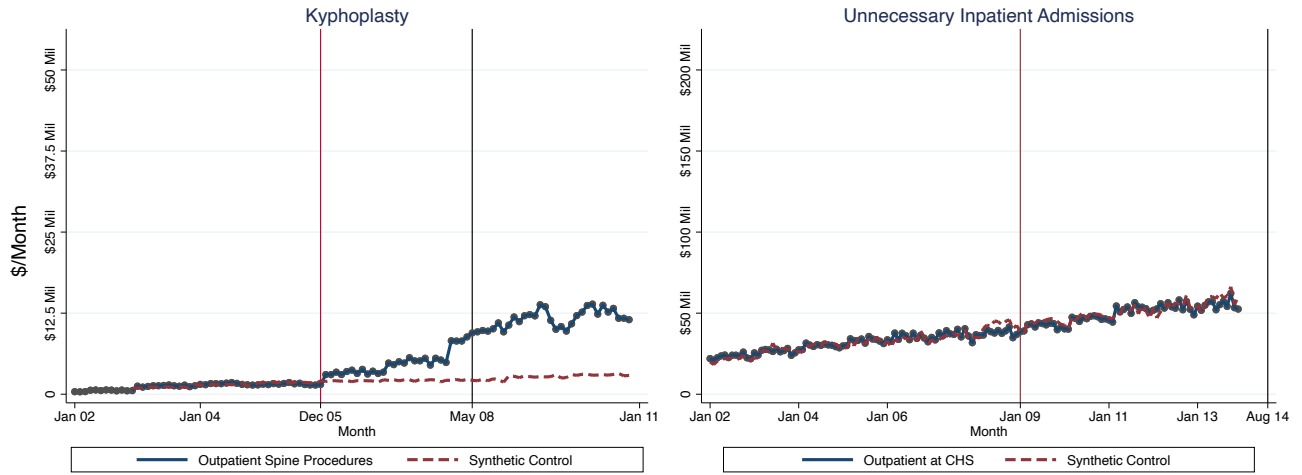
Notes: This figure plots the substitution effect to outpatient spending for the Kyphoplasty (left) and Unnecessary Inpatient Admission (right) case studies. These graphs correspond to the bottom half of Figure 3 and are scaled identically to those panels for comparison. In both case studies, whistleblowers alleged that patients should have been treated outpatient instead of inpatient. Outpatient spine procedure spending (left) rose following the kyphoplasty case as compared to the synthetic controls. However, there is no increase in outpatient spending at defendant hospitals (right) following the unnecessary admissions case. For the Unnecessary Inpatient Admissions case, multiple defendant hospitals were analyzed, and the series included here reflects Community Health Systems, the largest defendant hospital chain. Appendix Figure A3 plots the same figure for the other defendants in that case.

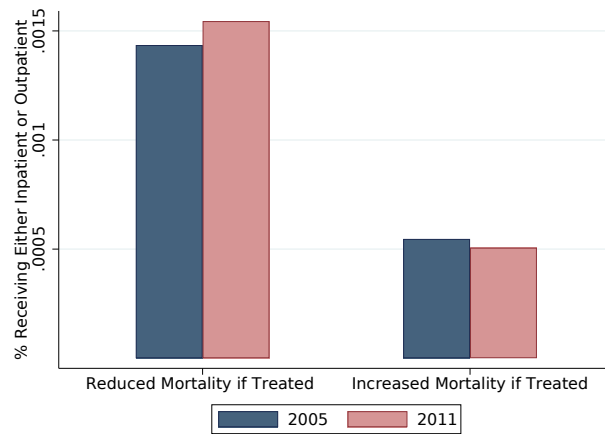Figure 5: Kyphoplasty Inpatient Treatment by Heterogeneous Treatment Effect
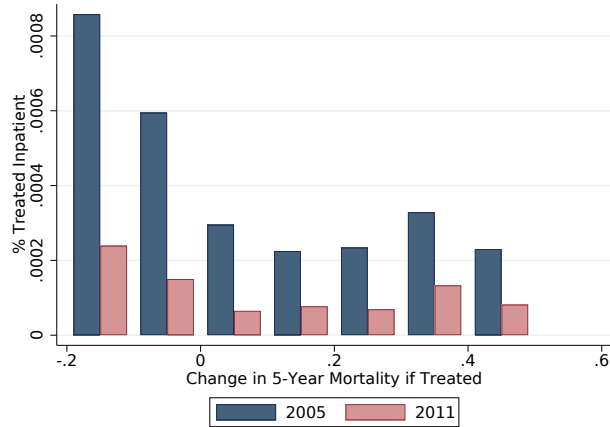
Notes: This figure plots the probability of receiving short stay inpatient kyphoplasty among the 2005 and 2011 cohorts, by the predicted treatment effect. Treatment effect is scaled as the difference in the probability of death in the next 6 years if one receives inpatient treatment, and negative values correspond to a lower probability of dying. Absolute treatment probabilities are low, reflecting the inclusion of the full population in this analysis and the relative rarity of kyphoplasty. In both cohorts, patients with higher expected benefits are more likely to receive the treatment. The reduction in treatment probability occurs evenly across the treatment effect distribution.

Figure 6: Outpatient Treatment by Heterogeneous Treatment Effects
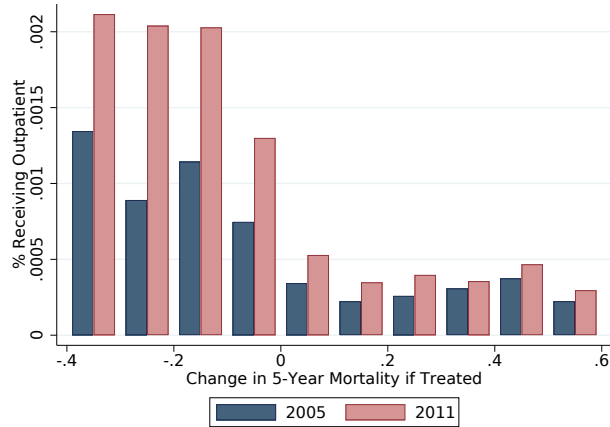
Notes: This figure plots the probability of receiving outpatient kyphoplasty or vertebroplasty by expected treatment effect among the 2005 and 2011 cohorts. The treatment effect is scaled as the change in probability of death in the next 6 years if one receives inpatient treatment; negative values indicate a lower probability of dying if treated. The whistleblower lawsuit settled in 2008 alleged that patients should have been treated outpatient instead of inpatient, and correspondingly, patients in 2011 were much more likely to receive outpatient treatment. These gains are greatest among patients to the left of the treatment effect distribution, which corresponds to the greatest benefits from the procedure.

Figure 7: Inpatient + Outpatient Treatment Probability by Health Benefit

Notes: This figure plots the probability of receiving either inpatient or substitute outpatient treatment by the predicted inpatient treatment effect, before and after the 2008 whistleblower settlement concerning kyphoplasty. Patients in 2011 who are expected to benefit from the procedure were 7% more likely to be treated in 2011 than in 2005, and patients who are expected to be harmed were 7% less likely to be treated.

Table 1: Summary Statistics about Whistleblower Lawsuits

	Healthcare Whistleblower Lawsuits	All Whistleblower Lawsuits
Number of Lawsuits	3269	5967
Years	1986-2012	1986-2012
Government Intervened	31.6%	27.5%
Of Which Settled/Judged in Favor of Gvt.	91.0%	92.3%
Settled/Judged in Favor of Government	35.7%	32.2%
Of which Government Intervened:	80.5%	78.7%
Settlement Amounts (Among Cases Not Dismissed)		
Mean	$22,678,349	$17,450,907
Median	$1,500,000	$1,315,540
Standard Deviation	$87,029,658	$71,445,704
Total	$26.47 billion	$33.54 billion
Whistleblower Share (Among Cases Not Dismissed)		
Mean	$3,798,847	$2,968,751
Median	$250,614	$228,750
Standard Deviation	$14,891,989	$12,226,857
Total	$4.23 billion	$5.49 billion
Case Length (Days)		
Mean	1138	1140
Median	964	920
Standard Deviation	800	883

Notes: This table presents descriptive statistics about False Claims Act whistleblower lawsuits using data from a Freedom of Information Act Request filed with the Department of Justice.

Table 2: Case Studies of Medicare Whistleblowing Enforcement

Type of Care	Type of Fraud	First Case Filed	First Settlement	# DOJ Press Releases	Settlement Total
Inpatient	Manipulation of Outlier Payments	Nov, 2002	Dec, 2004	11	$923,033,623
Botox	Off-Label Promotion	June, 2007	Aug, 2010	1	$600,000,000
Kyphoplasty	Inpatient Procedure Should be Outpatient	Dec, 2005	May, 2008	10	$214,238,775
Inpatient	Unnecessary Hospital Admissions	Nov, 2004	Dec, 2007	7	$172,296,460

Notes: This table shows the 4 highest settlement value case studies of Medicare whistleblowing enforcement for which I have data. Case studies are constructed using Department of Justice press releases to link lawsuits with similar allegations. Case studies where the first lawsuit was filed before the start of the data are omitted, as are cases that concern allegations not potentially observable in the Medicare claims data. Appendix C contains more details about the grouping of lawsuits and on the potential case studies not analyzed here.

Table 3: Deterrence Effects of Major Whistleblowing Categories

Type of Care	Type of Fraud	Settlement Total	Specific Deterrence	Deterrence Ratio
Inpatient	Manipulation of Outlier Payments	$923 Million	$17.46 Billion	18.92
Botox	Off-Label Promotion	$600 Million	-$41.67 Million	- .069
Kyphoplasty	Inpatient Procedure Should be Outpatient	$214.2 Million	$281.1 Million	1.31
Inpatient	Unnecessary Hospital Admission	$172.3 Million	$1.221 Billion	7.09
	Total	**$1.91 Billion**	**$ 18.92 Billion**	
			Average Ratio	**6.81**

Notes: This table summarizes the results of case studies on the 4 largest categories of Medicare whistle-blowing enforcement. Specific deterrence values are computed using a staggered synthetic control strategy to compare treated units to their counterfactual in the absence of whistleblowing. The specific deterrence is computed over 5 years post-treatment with a 10% annual discount rate compounded monthly. The deterrence ratio is computed as the ratio of the deterrence value to the settlement total.

Table 4: Placebo Tests for Synthetic Controls

Case	Deterrence Value	1-Tail Placebo Test		
Outlier Payments	+$17.46 Billion	0.0 ($n = 5$)		
Botox	-$41.67 Million	0.03 ($n = 93$)		

Case	Inpatient Deterrence	1-Tail Placebo Test	Outpatient Deterrence	1-Tail Placebo Test
Kyphoplasty	+ $538.9 Mil	0.13 ($n = 30$)	-$257.8 Mil	0.067 ($n = 15$)
Unnecessary Inpatient Admission:				
Defendant: St Joseph's Atlanta	+$44.8 Mil	0.01 ($n = 100$)	-$27.4 Mil	0.00 ($n = 100$)
Defendant: Wheaton Hospital	+$5.8 Mil	0.13 ($n = 100$)	-$83.8k	0.31 ($n = 100$)
Defendant: El Centro Medical Center	+$5.3 Mil	0.35 ($n = 100$)	-$4.0 Mil	0.02 ($n = 100$)
Defendant: Overlook Hospital	-$16.0 mil	0.02 ($n = 100$)	+$10.7 mil	0.08 ($n = 100$)
Defendant: Morton Plant Hospitals	+$266.6 Mil	0.01 ($n = 100$)	+$12.7 Mil	0.07 ($n = 100$)
Defendant: Shands Hospitals	+ $124.2 Mil	0.02 ($n = 100$)	+$50.7 Mi	0.00 ($n = 100$)
Defendant: Community Health Systems	+$693.2 Mil	0.07 ($n = 100$)	+$54.5 Mi	0.20 ($n = 100$)

Notes: This table summarizes the placebo test for the synthetic control strategy. For each control group, I compute the placebo deterrence effect, using the staggered synthetic control method with all other controls. The 1-tail test counts how many placebo groups exceed the deterrence value of the treated unit. For the kyphoplasty and unnecessary admissions cases, this test is conducted separately for the inpatient and outpatient spending. Deterrence effects are positive if spending on the treated unit is less than the control unit, and negative if spending on the treated unit is greater than the control unit.

Appendices

A   Cleaning of the FOIA Data on Qui Tam Whistleblower Suits

Data on the full set of whistleblowing lawsuits were gathered from a Freedom of Information Act (FOIA) request I conducted on the Department of Justice. For each lawsuit, the available data include: the docket number, district of filing, and case caption; the date the Attorney General was served notice of the suit; the primary federal agency to which the lawsuit related; whether or not the government intervened, and what date that election was made; the date of the settlement, judgement, or dismissal; the settlement amount if any, and the whistleblower's share. Each line of the FOIA dataset contains information about a suit that was dismissed, or in the event of a settlement, a settlement value related to that suit. Lawsuits against multiple defendants can have more than one settlement, and therefore appear in more than one line of the data. To correct this issue, I collapse the data by docket, filing district, and year: if two lawsuits contain identical docket numbers and were filed within the same state and year, I assume they are a single suit, and create a total of their settlement values. For the descriptive statistics on medical-related lawsuits, I restrict to suits for which the primary federal agency is either Health and Human Services, the Center for Medicare and Medicaid Services, or the Food and Drug Administration.

B   Staggered Synthetic Control Methodology

B.1   Motivating Model

Consider spending on a type of medical care that could be affected by whistleblowing treatment. For $i = 1, \ldots, N$ and time t, we would observe the spending level Y_{it}^U in the absence of treatment and Y_{it}^I following treatment. Call T_i the treatment period for unit i, which is the filing of the whistleblower's lawsuit. Whistleblowing treatment has no effect on periods $t < T_i$; therefore $Y_{it}^U = Y_{it}^I$ for all $t < T_i$.

Let δ_{it} be the effect of treatment at time t. Because Y_{it} represents spending, δ_{1t} represents the change in spending on a procedure following whistleblowing. Thus the spending level can be written as:

$$Y_{it}^I = Y_{it}^U + \delta_{it} \mathbb{I}_{t \geq T_i}$$

Let $i = 1$ be the unit treated by whistleblowing, which is the only unit subject to treatment; thus $Y_{it}^U = Y_{it}^I$ for control units $i > 1$ for all t. The treatment effect of interest is δ_{1t}, which is given by:

$$\delta_{1t} = Y_{1t}^I - Y_{1t}^U$$

in periods $t \geq T_i$.

Because Y_{1t}^I is always observed for all times $t \geq T_i$, estimation of the treatment effect relies on estimation of Y_{1t}^U, which is not observed in post-treatment periods.

Suppose that control units exhibit similar time trends at different points in calendar time, beginning at t_{0i}, which varies between units. Suppose that for all units, Y_{it}^U is given by a factor model, the same assumption of Abadie et al. (2010):

$$Y_{it}^U = \kappa_\tau + \lambda_\tau \mu_i + \varepsilon_{it} \tag{8}$$

Here, $\tau = t - t_{0i}$ is the time after the start of the control unit's trend begins; κ_τ is a common time effect across all units at time τ relative to the unit's start of the trend; λ_τ is a vector of common factors describing the trajectory of an outcome along a common trend; the parameter μ_i is an unknown vector describing the individual factor weights; and ε_{it} is a set of unobserved shocks of 0 mean.

Consider a $(N-1 \times 1)$ vector of weights $\vec{W} = (w_2, w_3, \ldots, w_N)$, such that $w_i \geq 0$ for $i = 2, \ldots, N$ and $\sum_{i=2}^N w_i = 1$. These values represent weights on the untreated control units, and every value of the vector $\vec{W}$ represents a possible synthetic control. Then, a weighted average of the control units is given by:

$$\sum_{i=2}^N w_i Y_{it} = \kappa_\tau \sum_{i=2}^N w_i + \lambda_\tau \sum_{i=2}^N w_i \mu_i + \sum_{i=2}^N w_i \varepsilon_{it}$$

If weights w_i^* can be constructed such that:

$$\sum_{i=2}^N w_i^* \mu_i = \mu_1$$

Then it holds that

$$E\left[\sum_{i=2}^N w_i^* Y_{it}\right] = E\left[\kappa_\tau + \lambda_\tau \sum_{i=2}^N w_i^* \mu_i + \sum_{i=2}^N w_i^* \varepsilon_{it}\right]$$

$$= E[\kappa_\tau + \lambda_\tau \mu_1] + \sum_{i=2}^N w_i E[\varepsilon_{it}] = E[Y_{1t}^U]$$

Therefore, the weighted average of the control units provides an unbiased estimator of the untreated counterfactual of the treated unit:

$$\sum_{i=2}^N w_i^* Y_{it} = \widehat{Y_{1t}^U}$$

In practice, these weights will be estimated, which means that there will be bias in the synthetic control estimate. Considerable other research has addressed the issues of this bias, such as Ferman and Pinto (2021) and Ben-Michael et al. (2020). I make the same argument as Abadie et al. (2010), that the bias of the estimator is bounded if the pre-treatment fit is good for a long series of pre-treatment periods.

Given these weights, we can estimate $\widehat{\delta_{1t}} = \widehat{Y_{1t}^U} - Y_{1t}^I$. Here, $\widehat{\delta_{1t}}$ is the change in spending due to whistleblowing. By integrating $\widehat{\delta_{1t}}$ over the post-whistleblowing periods, we can estimate a discounted specific deterrence effect:

$$D = \int_{t=T_1} (\widehat{Y_{1t}^U} - Y_{1t}^I) \beta^{t-T_1} dt \tag{9}$$

where T_1 is the treatment period and β^{t-T_1} is a discount factor starting at the treatment period.

B.2 Implementation

The practical estimation of the model presented in Appendix B.1 can be performed as a two-step procedure: estimating the time shift for each control unit, and then finding synthetic control weights w_i. Figure 2 provides a simple graphical explanation of the time-shifting process for two controls, one shifted forward in time and one shifted backward.

First, I consider the set of control units on similar pre-treatment trends to the treated unit. These units are qualitatively similar, representing similar types of medical care. Inpatient Diagnosis Related Codes (DRGs) affected by whistleblowing are compared to other DRGs; hospitals are compared to other hospitals. This follows directly from the suggestion in Abadie et al. (2010), who state that "researchers trying to minimize biases caused by interpolating across [units] with very different characteristics may restrict the donor pool to [units] with similar characteristics to the [unit] exposed to the event or intervention of interest."[19] Control units are chosen to be on similar pre-treatment trends to the treated unit; in general, they are on similar upward trajectories to the treated unit in the pre-treatment period.

I align the control units with the treated unit to ensure they are experiencing common time trends in the pre-treatment period. For each control, I construct a set of leads and lags, and find the lead or lag with the best fit to the treated series in the pre-period. With any fixed set of data, producing leads and lags creates missing data at the front or back of the series: in a monthly series, if one uses a 5 month lag, the first 5 months of available data have no value. In practice, this means that shifting the control units too far forward or back in time leaves a limited set of data for the evaluation of pre-treatment fit and post-treatment effects. Here, I bound the time shifts to ensure that there are 36 months of pre-treatment data, used to construct the synthetic control weights, and 60 months of post-treatment data, used to compute the deterrence effect.[20] Within these bounds, I select the appropriate lead or lag for each control unit that minimizes average square distance from the treated unit in the pre-period:

$$\min_d \frac{\sum_t^M (Y_{1t} - Y_{it+d})^2}{M} \tag{10}$$

for control unit i, where d indexes the different leads and lags, and M is the number of pre-treatment periods in which the shifted control and the treated unit overlap.[21] The use of a single lead or lag for each potential control reduces the dimensionality of the donor pool when fitting the synthetic control weights, mitigating concerns about overfitting that can arise from having too many controls (see, e.g. Ferman (2020)).

After associating each control with a time shift, I conduct the standard synthetic control process to choose weights as per Abadie et al. (2010). Weights are chosen to minimize mean-square error over all pre-treatment periods in which all of the controls overlap:

$$\min_{\vec{W}} \sum_{t \in M^*} \left(Y_{1t} - \sum_{i=2}^N w_i Y_{it+d_i^*} \right)^2 \tag{11}$$

where M^* is the set of periods for which all of the time-shifted controls overlap with the treated unit; d^* is the optimal time shift found by 10; and $\vec{W}$ is the set of all potential $(N-1 \times 1)$ vectors of weights $(w_2, \ldots, w_N)$ where $w_i \geq 0$ and $\sum_{i=2}^N w_i = 1$. Given the optimal time shift for each control, the Stata package "Synth" finds the optimal

---

[19] Abadie et al. use the word "region", not "unit" in this quote, but the authors also directly indicate that "region" is a generic term. They write: "we adopt the terms 'region' or 'unit' and 'intervention' or 'treatment,' which can be substituted for 'country,' 'state,' 'city,' etc. and 'event,' 'shock,' 'law,' etc., respectively for specific applications."

[20] In circumstances where the treatment period is too close to the start of the data, 36 pre-treatment periods are unavailable, and all of the available periods are used.

[21] Average square distance and not total square distance must be used because the number of points over which this sum is evaluated depends on the time shift.

weights w_i^*.[22]

Once these weights are found, the synthetic control unit is produced as the weighted sum of the control groups:

$$\widehat{Y_{1t}^U} = \sum_{i=2}^{N} w_i^* Y_{it+d^*}$$

Appendix Table A1 presents the root mean square prediction error (RMSPE) of the pre-treatment fit of the synthetic control group on the treated unit. The small RMSPE values for each case validate the assumption of good pre-period fit necessary to ensure that the bias of the synthetic control method is small.

In this paper, the outcome variables Y_{it} are all spending amounts. Therefore, the difference between the synthetic control and treated unit is a difference in spending, which can be integrated over the post-treatment periods. I estimate the specific deterrence effect as sum of the discounted difference between the treated and synthetic control over 5 years post-treatment:

$$\widehat{D} = \sum_{t=T}^{T+60} \frac{\widehat{Y_{1t}^U} - Y_{1t}}{(1.1^{1/12})^{t-T}} \tag{12}$$

where t is the time in months, T is the treatment period, and the denominator $1.1^{1/12}$ provides the monthly rate for a 10% annual discount rate. Deterrence is totaled for 5 years from the treatment date of filing. A positive deterrence value indicates that post-whistleblowing spending Y_{1t} is lower than that of the synthetic control group $\widehat{Y_{1t}^U}$.

There are some circumstances where the reduction in spending due to whistleblowing is immediately offset by an increase in spending on a substitute procedure. For example, in two of the whistleblower lawsuits described below, the allegations centered around the unnecessary use of inpatient medical procedures that could have been performed outpatient at lower cost. In these circumstances, I estimate the treatment effect for both the main (inpatient) and substitute (outpatient) effects and compute deterrence as the net reduction in spending:

$$\hat{D}_{net} = \hat{D}_{main} - \hat{D}_{substitute} \tag{13}$$

C Constructing Case Studies from DOJ Press Releases

The FOIA data described in Appendix A present a complete set of court-related information, but do not give information about the alleged behaviors for which the whistleblower sued, which are necessary for the case studies conducted here. To find details about the nature of these lawsuits, I scraped the Department of Justice press release archives for all press releases that contain the words "false claims," "Medicare," and either "qui tam" or "whistleblower." The DOJ makes an effort to publicize all of its successful cases, in particular because this strengthens later cases against providers who claim ignorance about what conduct constitutes an FCA violation.

From this universe of press releases, I group lawsuits with similar types of conduct into case studies. First, I read and hand-coded all press releases through 2014, which contained 325 press-releases. The majority of press releases describe settlements; however, press releases occasionally describe government intervention in a case, or provide year-end totals of successful recoveries, which I discarded. Then, each settlement press release was coded for the type of

---

[22]The two-step procedure for staggered synthetic controls is a tractable way to implement this methodology by leveraging existing methods. Separating the time shift component from the weighting component also ensures that synthetic control units are not constructed of multiple instances of the same control at different points in time.

medical care and the type of fraud it pertains to.

Certain types of care and certain types of fraud cannot be analyzed with my data and were omitted from the pool of potential cases to study. For example, cases regarding hospital cost reports, cases against Medicare claims processors, or cases that primarily concerned Medicare Advantage plans were discarded due to a lack of data. Similarly, some of the alleged frauds involve illegal kickbacks or improper financial relationships between providers. My available Medicare data do not contain financial structure information about providers, and so these types of cases were excluded from this study.

Following these restrictions, there are 170 remaining press releases that I group into potential case studies. Press releases are grouped by the type of fraud and the type of care they describe, and within each case study I create a total settlement amount. For 3 of the largest individual settlements, each against hospitals, the settlement press releases describe multiple types of allegations relating to different types of fraud reflected in other press releases. For these lawsuits, the settlements were apportioned to the different case studies based on how much money was paid for each type of conduct, as described in the settlement agreement or press release. For example, the June 2006 Tenet Healthcare settlement (described in Appendix D.4) was a $900 million settlement, but the press release states that only $788 million was for outlier payments while $46 million was for DRG upcoding. The outlier payment case study therefore is apportioned $788 million from this press release and the DRG upcoding case study gets $46 million.

This categorization process results in 54 distinct case studies. There are 23 cases with total settlements of less than $10 million and each contain 1 or 2 press releases. The top 11 case studies detail more than $100 million in settlements each; these cases are described in Appendix Table A3. If a lawsuit began before the data are available, I am unable to observe a pre-whistleblowing period, and therefore the case is omitted. In one case study, hospice care, there is insufficient data in the court filings or within the public records to identify the defendant providers, and this case is also omitted. Appendix Table A3 details the exclusion reasons for each of the top cases that were omitted, usually the timing of the first lawsuit. Researchers with access to earlier data may be able to conduct similar analyses on these examples of whistleblowing.

The press release data do not contain sufficient detail to conduct analyses in the Medicare data, only to generally compare allegations. To augment the details of the press releases, I collect whistleblower complaints and settlement agreements from the lawsuits detailed by the press releases. The identification of these cases is done either by docket number, which the press release sometimes specifies, or by defendant name. The FOIA data described in Appendix Section A were also used for mapping from press releases to court case docket numbers, which allowed for the retrieval of court documents. Whistleblower allegations and settlement documents contain specifics on the allegations of fraud or misconduct, including information on the medical coding of related procedures.

D   Lawsuit Details for Case Studies

D.1   Outlier Payment Case Study Details

Medicare reimburses most inpatient stays under a prospective payment system, with each stay classified under a Diagnosis Related Group (DRG). Hospitals are paid a fixed reimbursement for each DRG based on the average costs of treating patients under that DRG. This incentivizes providers to keep costs down, as they can recover profits by spending less per patient than the DRG pays. However, this contains the potential incentive to avoid treating high-cost patients. To correct this issue, Medicare has a system by which hospitals treating exceptionally high cost patients

receive additional reimbursements called outlier payments. The gravamen of the accusations in the outlier payment lawsuits were that the defendants manipulated the reimbursement process for outlier payments to classify more patients as outliers and receive additional payments.

Between December, 2004 and March, 2010 the Department of Justice published 11 press releases detailing settlements related to outlier payment falsification. The outlier-related conduct from these press releases totals to $923 million in settlements. The first press release for this case study was in December 2004, for the case US ex rel James Devage et al. v. HealthSouth et al. This lawsuit was originally filed in 1998; however, looking at the court documents from this case, whistleblowing was only a portion of this settlement, and the allegation of outlier falsification was not alleged by the whistleblower. Rather, it appears the Department of Justice included a provision for outlier falsification in this settlement at a later date. The first whistleblower complaint alleging outlier falsification comes from US ex rel. [Under Seal] v. Tenet Healthcare Corporation. et al., Case No. 02-8309, (E.D. Pa.). The filing of the Tenet Case, November 4, 2002, is used as the treatment date for this case. This lawsuit settled in June 2006 and was followed immediately by a Department of Justice press release. The Tenet settlement contains $788 million of recovery for outlier falsification, the bulk of the settlement total for this case study.

Outlier data were gathered from the 100% Medpar files, which detail each inpatient stay paid for by Medicare, from 1999–2016. There are more than 5 million total stays classified as cost outliers in this period, and at its peak usage in 2002 (pre-whistleblowing), outlier payments exceeded $500 million per month. The outlier payment system also theoretically contained a provision for outpatient outlier payments. However, in practice there are almost no outlier payments listed in the outpatient claims files, even at the height of inpatient outlier spending. This analysis is therefore restricted to inpatient cost outliers.

The control groups for the Outlier payment case are other types of expenditure that are of similar size and nature to outlier payments. Medicare pays for durable medical equipment (DME), home health aide services (HHA), hospice care (HOS), and skilled nursing facilities (SNF) as part of its broader package of benefits for older Americans. Spending on each of these types of care are included in the pool of potential controls. Furthermore, Medicare has a system for compensating hospitals that provide services to primarily low income patients, called disproportionate share hospital (DSH) adjustments. Much like outlier payments, DSH payments are an adjustment above regular inpatient DRG pricing.

Table A4 details the time shifts (in months) and synthetic control weights for these control groups in constructing a synthetic control unit. The synthetic control method places greatest weight to DSH payments, which are the most similar in nature to outlier payments and were also the subject of a later whistleblower lawsuit for improper use.

The time series of the Outlier payment expenditure shows a dip in outlier claims one month before the whistleblower filed. Inpatient stays claim processing takes time, and hospitals have up to one year to file a claim; in practice, they do so quickly to receive reimbursements, but not necessarily in the same month as the hospital stay. The whistleblower suit was filed during the first week of November of 2002, and we see a dip in October outlier claims, reflecting changes in billing practices for claims not yet filed at the time of the lawsuit.

D.2 Botox Case Study Details

The whistleblower lawsuits against Botox alleged that Botox was prescribed for non-FDA approved, non-Medicare-reimbursable uses. The whistleblowers further allege that Allergan, the maker of Botox, explicitly promoted the product for these "off-label" uses, giving Allergan civil liability for the False Claims made to the Medicare and Medicaid programs. In September 2010, Allergan settled with the Department of Justice to resolve 3 pending whistleblower

lawsuits of the same accusations: these cases have federal court docket numbers 1:07-cv-1288, 1:08-cv-1883, and 1:09-cv-3434, all conducted in the Northern District of Georgia. The first lawsuit was filed on June 5, 2007, which is used as the treatment date for this case. As part of this settlement, Botox agreed to pay $600 Million to the federal government, which includes both a civil settlement and a criminal penalty, for which whistleblowers received $37.8 million. This settlement was described in a Department of Justice press release in September, 2010.

Botox injections are outpatient procedures. Outpatient treatments are given a Current Procedural Terminology (CPT) or Healthcare Common Procedure Coding System (HCPCS) code that determines the reimbursement for the procedure, and an ICD-9 diagnosis code for the condition being treated. Documents from the whistleblower lawsuits provide details on the coding of outpatient Botox procedures. Medicare allowed reimbursement for Botox injections coded under CPT/HCPCS codes 64612, 64613, 64614, 64640, 64650, 67345, or J0585. The settlement agreement specifies that it resolves liability for false claims under ICD-9 diagnosis codes for spasm of muscle (728.85), other facial nerve disorders (351.8), spasmodic torticollis (333.83), unspecified torticollis (723.5), and bladder conditions (788.30 through 788.34, and 599.82).

Botox spending data were compiled from 100% samples of outpatient claims plus 20% samples of Carrier File (physician office visit) claims from January 2002-September 2015, using the CPT codes listed above and filtered for claims where the principal diagnosis matched the ICD-9 codes specified in the settlement. Data start at 2002 due to the availability of cleaned files, and data are truncated from October 2015 onwards due to the change from ICD-9 to ICD-10 diagnosis codes. To construct a full estimate of spending, spending on each outpatient CPT/HCPCS code from the 20% carrier file was multiplied by 5, then added to the spending from the 100% outpatient file. Spending for Botox under the relevant diagnoses codes grew from $20 million dollars in 2003 to $39 million in 2006, the year before the lawsuit against Allergan was filed.

There are 40,401 CPT/HCPCS codes observed in our data, motivating a restriction of these groups to better potential controls. The use of too many potential controls for the synthetic control method can result in overfitting (Ferman, 2020), so it is inappropriate to allow the synthetic control method to select controls from the entire pool. The candidate control groups used for this study are all other outpatient CPT/HCPCS codes for which spending started within 10% of the range of Botox's spending and saw a rise over any 3-year period between 2002 and 2011 within 25% of Botox's observed rise, of which there are 93 control units. Table A5 shows the weights and time shifts for the 10 control groups given the highest weights by the synthetic control method.

D.3 Kyphoplasty Case Study Details

The main allegations of the kyphoplasty lawsuits were that hospital, at the urging of the product manufacturer Kyphon, conducted kyphoplasty as an inpatient procedure rather than outpatient. Under Medicare, inpatient stays are paid a fixed amount for the Diagnosis Related Groups (DRG) under which a patient is coded. Therefore, for short inpatient stays, providers receive the full reimbursement and incur relatively low costs. Kyphon allegedly instructed its sales representatives and marketers to push usage of the DRGs 233, 234, and 216, which are various non-specific inpatient spine surgery codes not intended for kyphoplasty. The specific descriptors of these DRGs were, in 2005, the year the lawsuit was filed: DRG 234: "Other musculoskeletal system & connective tissue O.R. procedure without comorbidities and complications"; DRG 233, ibid, "... with comorbidities and complications"; and DRG 216: "biopsies of musculoskeletal system & connective tissue." (Center for Medicare and Medicaid Services, 2005) The whistleblowers further allege that Kyphon sold a bone biopsy kit that they encouraged physicians to use on all kyphoplasty patients to greater reimbursement through use of DRG 216.

Tracing spending on DRGs across time requires cross-walking when new versions of the DRG coding are released. This occurred once in the relevant time period, in October 2007. This change was a complete overhaul of the DRG system, and changed from DRGs to a severity-based system (called MS-DRGs). Under this change, sets of 1 to 2 DRGs before October 1, 2007 usually correspond to 3 DRGs after that date. No 1 to 1 crosswalk exists, and so I collapse the DRGs into groups which can be cross-walked through this change. The DRGs allegedly promoted by Kyphon exhibit this pattern: DRG 216 became MS-DRGs 477, 478, and 479, and DRGs 233-234 became MS-DRGs 515, 516, and 517. I create groups for the DRGs that map across this change, and these DRG groups provide the control units. I omit DRGs that were entirely eliminated or newly generated during this switchover, as they cannot be analyze across the relevant time period. There was a second DRG coding change in October 2015, but this change was close to the end of the available data and happened many years after the relevant lawsuit, so these are not necessary for analysis. Inpatient data from after October 2015 are dropped when constructing control units.

The treated unit for this analysis is the total payment for stays of 7 nights or fewer under the groups corresponding to DRGs 233, 234, and 216, the DRGs allegedly promoted by Kyphon. The set of controls are payments for stays of 7 nights or fewer under other DRG groups. I include DRG groups which experienced a more than double growth in annual spending over any 3-year period before 2011. The restriction to growing groups picks DRG groups on similar trajectories to the treated unit, which experienced a 2.5 times increase between 2002 and 2004, the year before the lawsuit was filed. The cutoff for growing controls is placed at 2011 to ensure that the data can be shifted back to match the kyphoplasty series and still allow for 5 years of post-treatment comparison, as my data end in 2015. I exclude DRGs which saw discontinuous jumps (a 500%+ increase in any single month, likely reflecting a major coding change rather than a usage change), or which were not in use for 12 or more months of the pre-whistleblowing period. There are 30 DRG groups included as controls. Appendix Table A6 details the time shift and synthetic control weights for these DRGs.

The kyphoplasty lawsuits alleged that kyphoplasty should have been coded as an outpatient procedure rather than inpatient. Outpatient procedures are billed to Medicare under HCPCS codes. Kyphoplasty was a new technology during this period, and coding for it changed over the course of the relevant period. Kyphoplasty was often billed under the catch-all unlisted spine procedure code 22899, but also was coded under the HCPCS codes 22523, 22524, 22525, 22513, 22514, 22515, C9718, or C9719 at various times, the latter two very infrequently. Furthermore, to measure substitution effects to outpatient procedures, I need to consider spending on vertebroplasty, a close substitute procedure, which was coded under HCPCS codes 22520, 22521, 22522, 22510, 22511, or 22512.

For the purposes of the health analysis in Section 7, the codes listed in the previous paragraph are used to identify outpatient kyphoplasty and vertebroplasty, as almost everything billed under these codes were in fact those procedures. However, whistleblowers also alleged that Kyphon, the maker of the kyphoplasty kit, also pushed providers to miscode the procedure under HCPCS codes 22327, 22325, 22328 for open reduction of thoracic or lumbar vertebrae. Kyphoplasty is not an open procedure, but is rather percutaneous. To analyze the sum of the fiscal effects, and to construct appropriate control groups, the outpatient deterrence analysis considers spending on **all** outpatient spine procedures, in the CPT code range 22010-22899. Some of these procedures were unaffected by whistleblowing, and therefore will difference out on average before and after the treatment period and will not bias the deterrence measurement. As controls, I consider other categories of surgical outpatient procedures on the musculoskeletal system, all of which are in the 20000-29999 range, of which the treated unit is a subset. These categories are constructed from the AAPC Coder code ranges (AAPC Coder, 2019) and include procedures like shoulder surgeries, hip surgeries, etc. and are not substitutes for the treated procedure. Two other codes in this range, CPT Codes 20000 and 20005, which correspond

to surgical drainage procedures, were also included; these codes were deprecated in 2019. Table A6 gives the time shifts and weights for these control units.

D.4 Unnecessary Inpatient Admission Case Study Details

When a patient visits a hospital, particularly for emergency services, physicians at that hospital make a decision on whether to admit the patient for an inpatient stay, which generally results in an overnight stay of at least one night. Instead of admitting patients, doctors have the ability to treat a patient outpatient, or to hold them for observation without admission. Inpatient admission receives greater reimbursement than outpatient or observational care. Under Medicare rules, inpatient stays are reserved for acute illnesses, and hospitals are expected to conduct utilization reviews to ensure that patients are admitted appropriately. The allegations in this case study are that the defendant hospitals improperly admitted Medicare patients because of the greater reimbursement provided.

Between 2007 and 2014, the Department of Justice issued press releases detailing 7 settlements with different providers and provider chains regarding this conduct. Four of the settlements concerned a single hospital: St Joseph's Atlanta; Wheaton Hospital in Wheaton, Minnesota; El Centro Medical Center in Southern California; and Overlook Medical Center in Summit, NJ. Two of the settlements concerned groups of 6 hospitals: Shands Hospitals and Morton Plant Hospitals, both in Florida. The final settlement was against Community Health Systems (CHS), described the Department of Justice in its press release as the "nation's largest operator of acute care hospitals." CHS settled for $98 million for conduct in 119 hospitals in 28 states. The total recovery from these 7 settlements was $172.29 million.

The evidence suggests that the conduct described in these cases was localized among the defendants. Appendix Figure A8 plots the total inpatient spending from all providers in the US and shows no changes with the filing of the first lawsuit in October 2004. This is unsurprising, as total Medicare inpatient spending was around $10 billion per month at the time of filing, and the total of these settlements was less than $200 million. Therefore, the computation of specific deterrence conducted here focuses only on the defendant hospitals. This may undercount spillover affects to other hospitals who were also deterred from unnecessary inpatient admissions as a result of these settlements.

The goal of this analysis is to measure the specific deterrence effects of these lawsuits on spending at the defendant providers. Because the lawsuits indicate that patients were unnecessarily admitted to the hospital rather than being seen outpatient, I expect a decrease in inpatient spending and an increase in outpatient spending. To measure this change, I construct control units using a set of untreated hospitals. Because some of the untreated hospitals may have been affected by spillovers, I restrict my control sample to hospitals in the 23 states (including the District of Columbia) with no defendant providers. These control units see different patient populations than the defendants and are less likely to be influenced by their behavior. This ensures the control units are isolated from the treated units, at least geographically, to mitigate spillover effects. Next, I construct a random sample of 100 control units for each defendant. For the four defendants that were 1 hospital, the control units are 100 randomly selected hospitals. For the two defendants which were 6 hospitals, the control units are 100 units of 6 randomly grouped hospitals, drawn with replacement from the set of control hospitals. For CHS, which had 119 hospitals settle, I construct 100 control units of 119 randomly grouped hospitals, drawn with replacement from the set of control hospitals. These control units serve as the controls for the inpatient spending. For outpatient spending, I repeat the same process, drawing from the set of outpatient providers in states with no defendants.

Each of the 7 defendants here is conducted as its own case study. Each has its own controls, and the treatment date for each defendant is the earliest filing date of the lawsuit(s) settled in the settlement agreement with that hospital. There are multiple lawsuits against some hospitals because of multiple whistleblowers. Because CHS constitutes 119

of the 135 hospitals in this study, plots from CHS are included in the main results. Inpatient and outpatient plots from the other defendants are presented in Appendix Figures A2 and A3 respectively.

Appendix Figures and Tables

Figure A1: Trends in Healthcare Whistleblowing Lawsuits
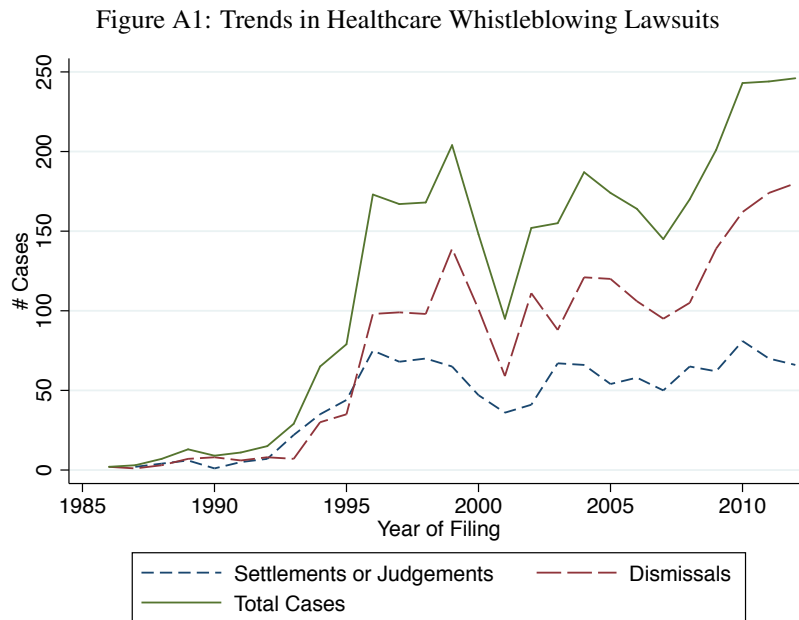
Notes: This figure plots the number of healthcare-related whistleblower lawsuits by year and splits the data by the outcome of the lawsuit. Data begin in 1986, when Congress amended the False Claims Act to allow for whistleblower lawsuits, and go through 2012, the last available year of data. Settlements rose to around 50 per year in 1995 and have stayed relatively constant, while total cases and dismissed cases have both continued to rise.

Figure A2: Inpatient Spending at Other Defendants in the Unnecessary Admissions Case Study
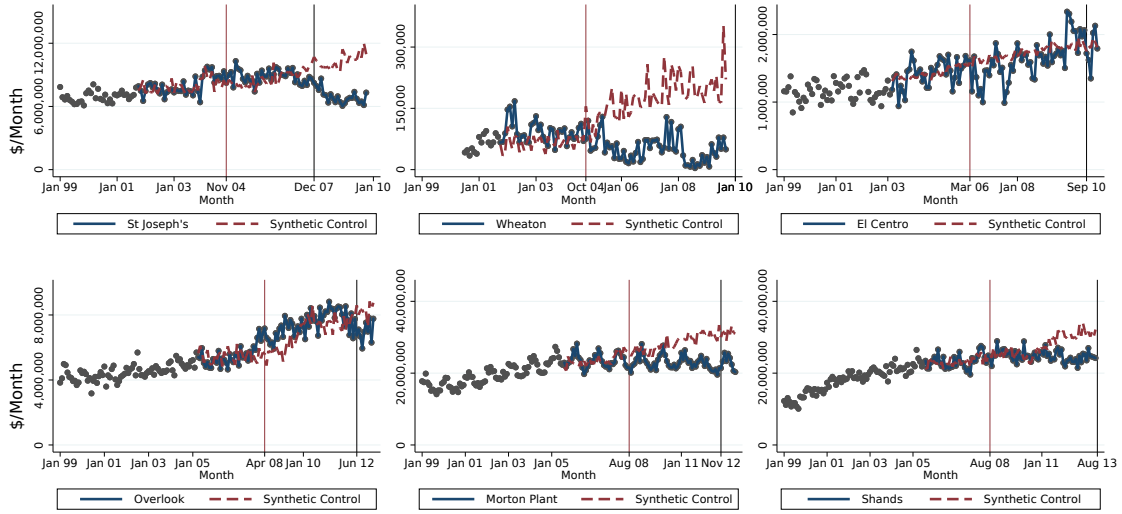

Notes: This figure plots the staggered synthetic control strategy for inpatient spending at the other defendant providers in the unnecessary inpatient admissions case. The largest defendant, CHS, appears in the bottom-right panel of Figure 3. On average, inpatient spending at these providers fell relative to the synthetic control group.

Figure A3: Substitute Outpatient Spending at Other Defendants in the Unnecessary Admissions Case Study
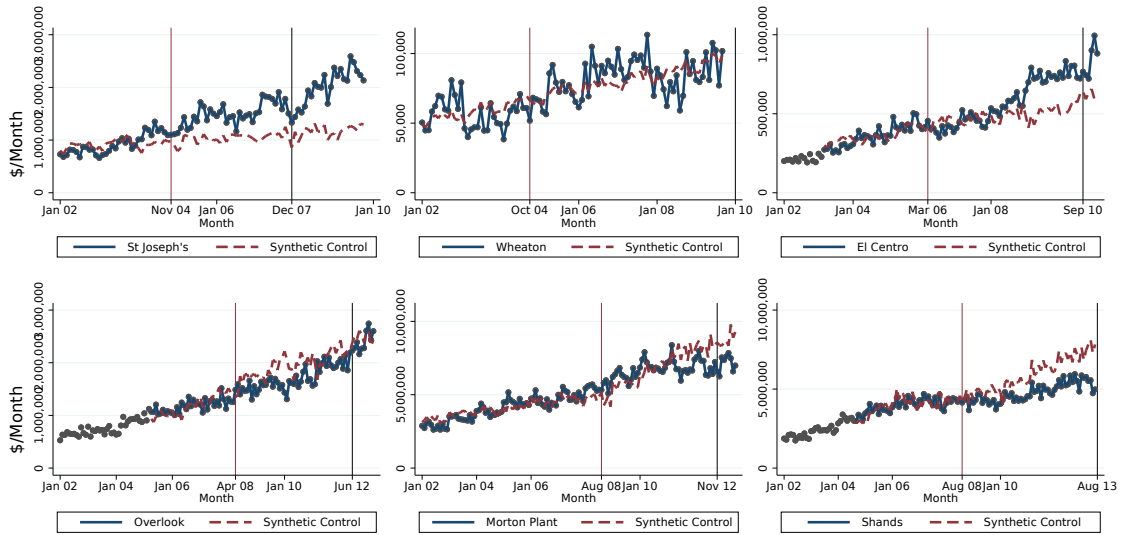

Notes: This figure plots the staggered synthetic control strategy for outpatient spending at the other defendant providers in the unnecessary inpatient admissions case. The largest defendant, CHS, appears in the right panel of Figure 4. On average, outpatient spending at these providers did not increase, even when inpatient spending fell. However, there is heterogeneity among the defendants, with some experiencing increases in outpatient spending and others experiencing decreases.

Figure A4: Robustness Check: Only Allowing Forward Time Shifts
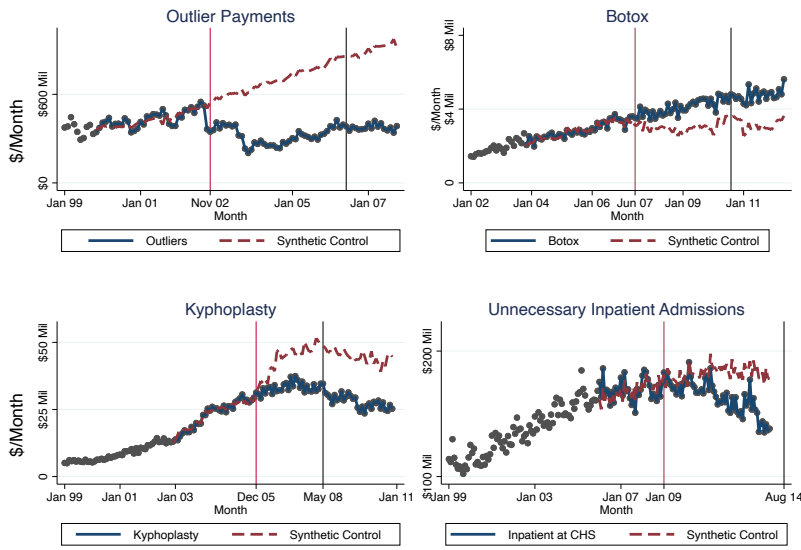
Notes: This figure repeats Figure 3, but only allows controls to be shifted forward in time. The results match the original specification.

Figure A5: Robustness Check: Removing Time Fixed Effects, Main Results
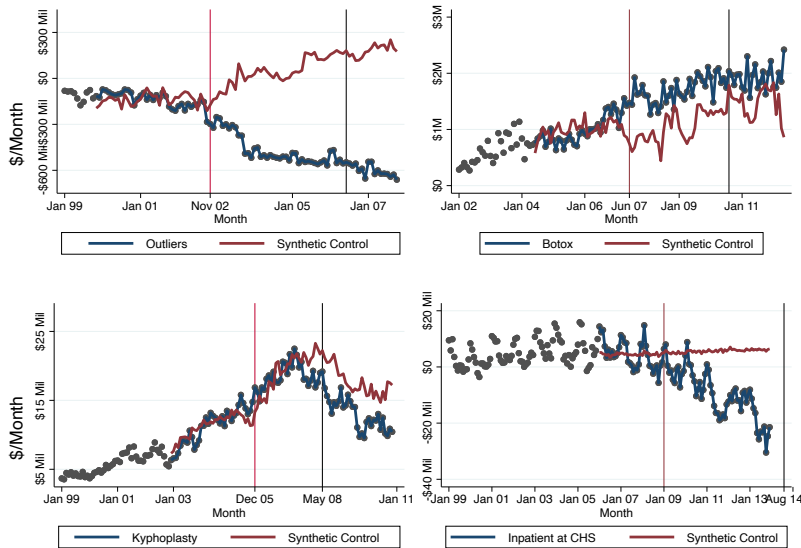
Notes: This figure repeats Figure 3, but removes fixed effects from the treated and control units before applying the synthetic control methodology. The results qualitatively match the original specification in terms of fit and directional trends, and the estimated total deterrence effects exceed the original specification.

Figure A6: Robustness Check: Removing Time Fixed Effects, Substitution Results
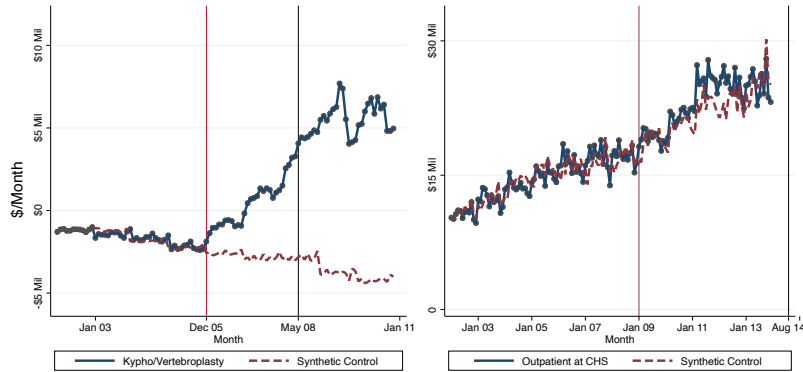

Notes: This figure repeats the substitution results from Figure 4, but removes fixed effects from the treated and control units before applying the synthetic control methodology. The results qualitatively match the original specification in terms of fit and directional trends, and the estimated total deterrence effects exceed the original specification.

Figure A7: Robustness Check: Flat Line Projection
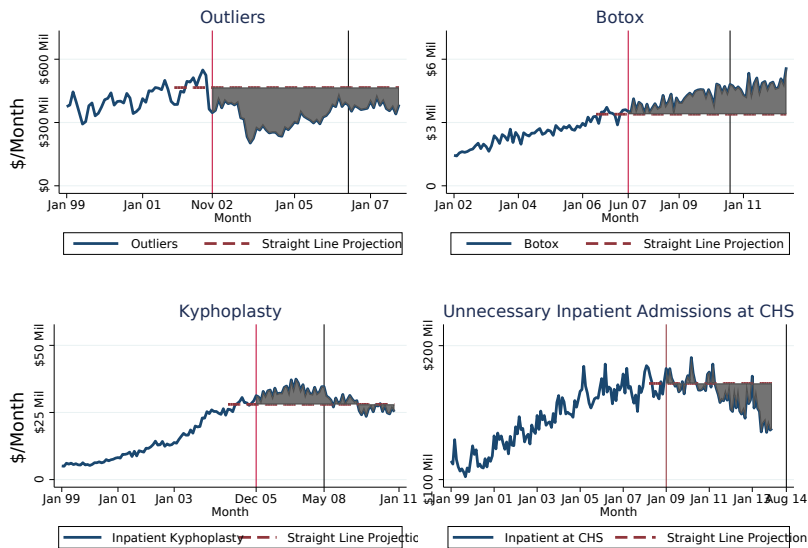
Notes: This figure repeats Figure 3, but replaces the control strategy with a flat line projection of the 12 months of spending prior to whistleblowing. The total deterrence measurement under this method, including substitution to outpatient spending for the Kyphoplasty and Unnecessary Inpatient Admissions Case, is $5.56 billion.

Figure A8: Total Inpatient Spending Over Time
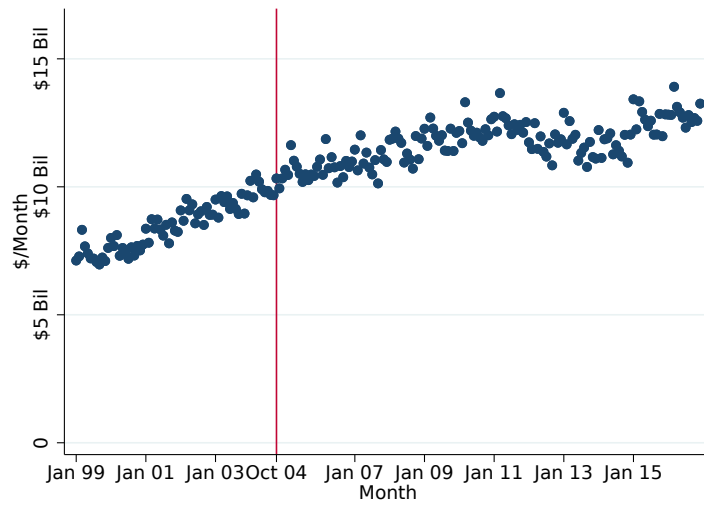
Notes: This figure plots total inpatient spending against the timing of the first unnecessary inpatient admissions lawsuit. There is no visible change in overall inpatient spending, which motivates an analysis focused on the defendants in these lawsuits.

Figure A9: One Night Inpatient Kyphoplasty Claims
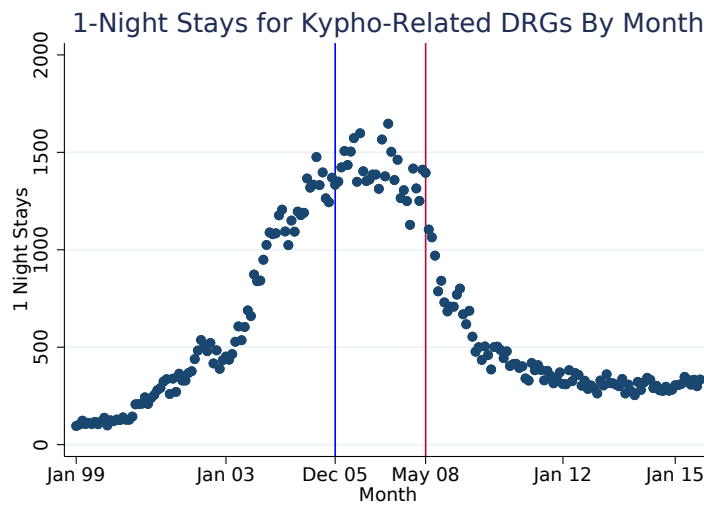
Notes: This figure plots inpatient stays for the DRGs promoted by Kyphon for inpatient kyphoplasty that lasted 1 night or less. The first vertical line shows the filing of the lawsuit, and the second line shows the settlement of the lawsuit.

Figure A10: Treatment Effect Histogram by Cohort
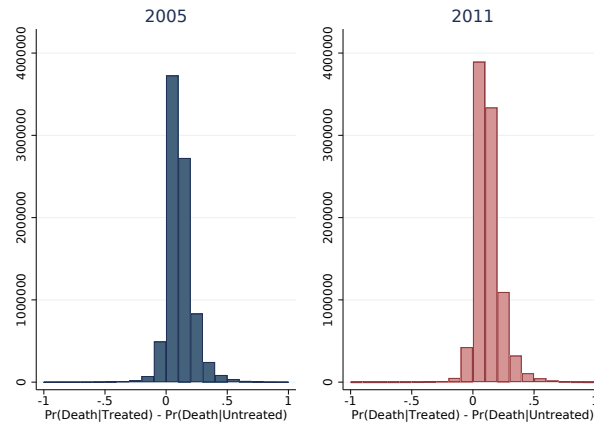
Notes: This figure plots the histogram of expected patient health effects from receiving a short-stay inpatient kyphoplasty treatment among the population of never-before-treated 70-75 year olds in 2005 and in 2011, which correspond to pre- and post-whistleblowing in the kyphoplasty case. Each cohort contains roughly 8 million patients. The horizontal axis is the difference in probability of death in the next 6 years if one receives treatment; values greater than 0 indicate a greater probability of death if treated, and negative values indicate a lower probability of death if treated. This treatment effect is computed using a model fit to the 2005 pre-whistleblowing cohort. The similarity of these histograms indicates that the sample population did not change in composition following the lawsuit.

Figure A11: Kyphoplasty Short-Stay Inpatient Treatment Count by Heterogeneous Treatment Effect
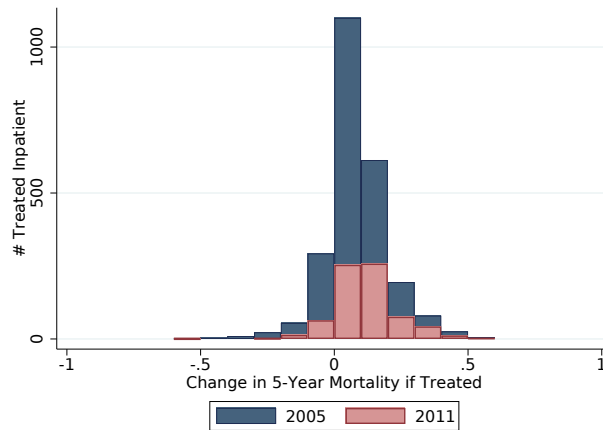
Notes: This figure plots the number of patients receiving short-stay inpatient kyphoplasty among the 2005 and 2011 cohorts of never-before-treated 70-75 year olds. Inpatient treatment counts were vastly reduced following the lawsuits against Kyphon and hospitals providing this treatment, which first settled in 2008. The treatment effect is identical to the horizontal axis in Figure A10, and is scaled as the change in probability of death when receiving treatment. The reduction in treatment volume occurs across the treatment effect distribution. The shape of these distributions is mostly driven by the number of units in each bin, as shown in Appendix Figure A10, motivating an analysis by probability of treatment as shown in Figure 5.

Figure A12: Inpatient or Outpatient Treatment Probability by Treatment Effect by Year
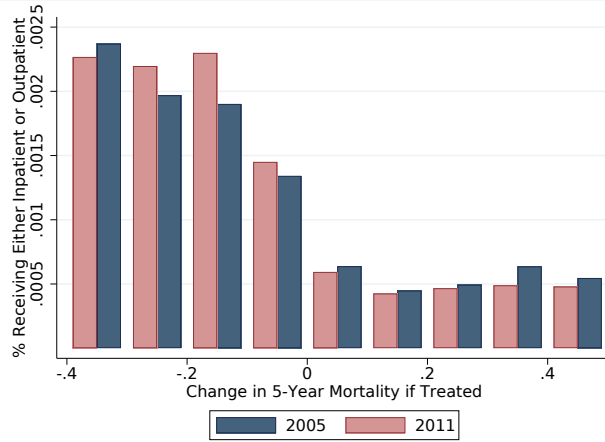
Notes: This figure plots the change in total (inpatient or outpatient) treatment probability as a function of the predicted treatment effect. It presents the same result as Figure 7, broken out by the treatment effect bin. Treatment effects are scaled as the change in probability of dying when receiving inpatient kyphoplasty. To satisfy Medicare cell-size-suppression rules, patients with treatment effects in the tails of the distribution are recoded to ± 0.4. Patients with beneficial treatment effects, i.e. less than 0, are on average more likely to receive treatment after whistleblowing, while patients that are expected to be harmed are less likely to receive treatment after whistleblowing.

Table A1: RMSPE of Synthetic Control Results

Case Study	Dependent Variable Pre-Period Mean	Synthetic Control RMSPE	Fraction RMSPE
Outlier Payments	$4.31*10^8$	$4.03*10^7$	0.0935
Botox	$2.72*10^6$	$2.10*10^5$	0.0773
Kyphoplasty – Inpatient	$2.29*10^7$	$1.90*10^6$	0.0830
Kyhoplasty – Outpatient	$1.45*10^6$	$2.23*10^5$	0.154
Unnecessary Inpatient Admissions – Inpatient at CHS	$1.69*10^8$	$7.03*10^6$	0.0416
Unnecessary Inpatient Admissions– Outpatient at CHS	$3.17*10^7$	$2.56*10^6$	0.0808

Notes: This table presents the Root Mean Square Prediction Error (RMSPE) of the staggered synthetic control strategy that estimates the main deterrence effect and substitution effect presented in Figures 3 and 4.

Table A2: Selected Logit Regression Coefficients for Heterogeneous Treatment Effects of Kyphoplasty

	Coef	SE	95% CI
Treated	2.876	2.863	[-2.736, 8.488]
Age	.0928	.000716	[.0914, .0942]
Treated $\times$ Age	-0.0117	0.0373	[-0.0848, 0.0614]
Female	-0.419	0.00222	[-0.423, -0.415]
Treated $\times$ Female	-0.193	0.139	[-0.465, 0.0782]
Race White	0.00245	0.0354	[-0.0670, 0.0719]
Treated $\times$ Race White	-1.35	0.679	[-2.68, -0.0220]
Race Black	0.0706	0.0356	[0.000847, 0.140]
Treated $\times$ Race Black	-1.32	0.811	[-2.91, 0.273]
OREC = DIB	0.526	0.00310	[0.520, 0.532]
Treated $\times$ OREC = DIB	-0.447	0.170	[-0.780, -0.116]
OREC = ESRD	0.558	0.0505	[0.459, 0.657]
Treated $\times$ OREC= ESRD	1.28	1.99	[-2.62, 5.19]
Previous Inpatient Stay	0.253	0.00283	[0.248, 0.259]
Treated $\times$ Previous Stay	-0.0688	0.134	[-0.331, 0.194]
Constant	-8.43	.0631	[-8.55, -8.31]

Notes: This table presents selected coefficients from the heterogeneous treatment effects regression described in Equation 5. The full model contains hundreds of coefficients due to the inclusion of state fixed effects and counts for stays under each inpatient DRG as well as full interaction with the treatment indicator. The coefficients presented here are given as examples. OREC indicates the original reason for Medicare qualification. ESRD denotes End Stage Renal Disease and DIB denotes disability insurance benefits.

Table A3: Potential Case Studies of Medicare Whistleblowing Enforcement

Type of Care	Type of Fraud	First Settlement Year	Settlement Total	Included or Omitted	Reason for Omission
Pharmaceuticals	Off-Label Promotion	2004	14,359,380,000	Omitted	Part D Data Start 2006
Inpatient	Outlier Payment Falsification	2004	923,033,623	Included	
Botox	Off-Label Promotion	2010	600,000,000	Included	
Inpatient	DRG Upcoding	2000	458,260,000	Omitted	Lawsuit Filed 1995
Home Health	Medically Unnecessary Care	2000	424,700,000	Omitted	Lawsuit Filed 1995
Nursing Home	Inadequate Care	2001	219,000,000	Omitted	Lawsuit Filed 1996
Kyphoplasty	Inpatient Should be Outpatient	2008	214,238,775	Included	
Physical Therapy	Unlicensed providers; Group Therapy Billed as One-on-One	2004	185,600,000	Omitted	Lawsuit Filed 1998
Hospital	Unnecessary Admissions	2007	172,296,460	Included	
Nursing Home Therapy	Falsified Hours Spent	2000	132,700,000	Omitted	Lawsuit Filed 1996
Hospice	Ineligible Patients	2006	114,886,000	Omitted	Defendants Not Identifiable from Court Data
Laboratory Tests	Medically Unnecessary; Unbundling Tests	1997	111,161,000	Omitted	Lawsuit Settled Before Data Start

Notes: This table describes the potential case studies of whistleblowing enforcement described in Appendix C. Each case study is constructed from a group of lawsuits. These are all of the case studies for which settlements totaled to more than $100 million. Four of the top case studies are conducted in this paper. Seven case studies are omitted because the first lawsuit was filed before the data are available. My available data start in 1999 for all types of Medicare except outpatient care and pharmaceuticals, which start in 2002 and 2006 respectively. One case study, ineligible hospice patients, is omitted because the lawsuit documents do not identify the defendant providers.

Table A4: Synthetic Control Weights and Time Shifts for Outlier Payments Case

Control	Time Shift (Months)	Synthetic Control Weight
DME	+9	.049
DSH	+1	.837
HHA	+10	.024
HOS	-23	.083
SNF	+10	.007

Notes: This table details the synthetic control time shifts and weights used for the Kyphoplasty case. The control units are other types of Medicare spending, described in detail in Appendix D.1. The time shift describes the number of months the control unit must be shifted to align with the treated unit in the pre-whistleblowing period. Positive values mean the control unit is shifted forward in time, and negative months mean the control is shifted back in time. For example, a time shift of +9 means that the control unit in March, 2005 serves as a control for the treated unit in December, 2005.

Table A5: Synthetic Control Weights and Time Shifts for Botox Case

CPT Code	Shortened Descriptor	Time Shift (Months)	Synthetic Control Weight
76775	Ultrasound, retroperitoneal	0	0.602
11043	Debridement, muscle and/or fascia	0	0.226
76830	Ultrasound, transvaginal (non-hyphenobstetrical)	+23	0.004
00300	Anesthesia (integumentary system,muscles and nerves of head, neck and posterior trunk), NOS	+40	0.003
01480	Anesthesia, open procedures on bones of lower leg, ankle, and foot, NOS	+40	0.003
14041	Adjacent tissue transfer or rearrangement	+37	0.003
22614	Arthrodesis, posterior or posterolateral technique, single level; each additional vertebral segment	+34	0.003
29580	Paste/Unna boot	+23	0.003
36245	Selective catheter placement, arterial system	0	0.003
43249	Esophagogastroduodenoscopy, flexible, transoral	+17	0.003

Notes: This table details the synthetic control time shifts and weights used for the Botox case. The control units are other types of outpatient care, described in detail in Appendix D.2. The time shift describes the number of months the control unit must be shifted to align with the treated unit in the pre-whistleblowing period. Positive values mean the control unit is shifted forward in time, and negative months mean the control is shifted back in time. For example, a time shift of +17 means that the control unit in June, 2005 serves as a control for the treated unit in November, 2006.

Table A6: Synthetic Control Weights and Time Shifts for Kyphoplasty Case

Inpatient

DRG V-24	MS-DRG V-25	Descriptor	Time Shift (Months)	Synthetic Control Weight
462	945, 946	Rehabilitation	47	0.431
533, 534	037, 038, 039	Extracranial Procedures	3	0.049
524	69	Transient Ischemia	5	0.045
518	250, 251	Percutaneous cardio procedures w/o coronary artery stent	17	0.045
535	222, 223	Cardiac defibrilator implant with cardiac catheterization	-5	0.037
519, 520	471, 472, 473	Cervical spinal fusion	-12	0.035
155, 156, 567, 568	326, 327, 328	Stomach, esophagealm and duodenal procedures	-58	0.03
515	226, 227	Cardiac defibrillator implant w/o cardiac catheterization	21	0.029
523	896, 897	Alcohol/drug abuse or dependence w/o rehabilitation therapy	-58	0.026
496	453, 454, 455	Combined anterior/posterior spinal fusion	-58	0.025

Outpatient

CPT Code Range	Surgical Category	Time Shift(Months)	Weight
20000, 20005	Incision and Drainage	0	0.158
22900-22999	Abdomen	0	0.115
21920-21936	Back or Flank	0	0.096
21501-21899	Neck or Thorax	0	0.079
26990-27299	Pelvis or Hip	0	0.077
21010-21499	Head	0	0.076
27301-27599	Femur or Knee	0	0.061
27600-27899	Leg or Ankle	0	0.058
29000-29799	Casts	11	0.051
25000-25999	Forearm or Wrist	11	0.048

Notes: This table details the synthetic control time shifts and weights used for the Kyphoplasty case. The top panel describes the controls for inpatient spending, which are groups of other inpatient DRGs. The bottom panel describes the controls for outpatient spending, which are other CPT code ranges of surgery on the musculoskeletal system. These controls are described in detail in Appendix D.3. The time shift describes the number of months the control unit must be shifted to align with the treated unit in the pre-whistleblowing period. Positive values mean the control unit is shifted forward in time, and negative months mean the control is shifted back in time. For example, a time shift of +3 means that the control unit in September, 2005 serves as a control for the treated unit in December, 2005.