Ambulance Taxis: The Impact of Regulation and Litigation on Health-Care Fraud

Paul Eliason

University of Utah and National Bureau of Economic Research

Riley League

University of Illinois Urbana-Champaign

Jetson Leder-Luis

Boston University and National Bureau of Economic Research

Ryan C. McDevitt

Duke University and National Bureau of Economic Research

James W. Roberts

Duke University and National Bureau of Economic Research

We study the effectiveness of pay-and-chase lawsuits and up-front regulations for combating health-care fraud. Between 2003 and 2017, Medicare spent $7.7 billion on 37.5 million regularly scheduled ambulance rides for patients traveling to and from dialysis facilities even though many did not satisfy Medicare's criteria for receiving reimbursements. Using an identification strategy based on the staggered timing of regulations and lawsuits across the United States, we find that adding a

We thank Zarek Brot-Goldberg, Josh Gottlieb, Tal Gross, Jonathan Gruber, Atul Gupta, Ben Handel, Steven Medema, Matt Notowidigdo, David Ridley, Andrei Shleifer, Jonathan

Electronically published March 19, 2025

Journal of Political Economy, volume 133, number 5, May 2025.

https://doi.org/10.1086/734134

prior authorization requirement for ambulance reimbursements reduced spending much more than pursuing criminal and civil litigation did on their own. We find no evidence that prior authorization affected patients' health.

I. Introduction

Fraud poses a serious problem for Medicare: it both distorts patient care and wastes limited public resources. In 2019, improper payments that did not meet Medicare's statutory, regulatory, administrative, or other legally applicable requirements totaled $31.2 billion, or 7.4% of overall spending (CMS 2023). To combat and deter this fraud, the federal government uses two main approaches: litigation through the courts, which attempts to recover funds that have already been paid out, and administrative regulations, such as prior authorization, that prevent improper payments from being made in the first place. Although, in theory, both approaches can be used effectively, the costly and expansive monitoring required to implement wide-reaching regulations has prompted a long literature in law and economics favoring the use of targeted litigation instead (Coase 1960; Becker 1968). The enforcement of most US health-care policies reflects this view (OIG 2021), yet no large-scale empirical studies have compared the effectiveness of commonly used pay-and-chase litigation to preemptive regulations like prior authorization now being used extensively throughout the US health-care system.

In this paper, we study the unnecessary use of ambulances to transport patients between their homes and dialysis facilities to provide the first systematic empirical evidence that administrative regulations can reduce health-care fraud more effectively than relying solely on ex post litigation. Although Medicare reimburses ambulance rides for those with a demonstrated medical need for assistance, unscrupulous companies have exploited a historically lax enforcement of the rules to provide fraudulent

Skinner, Becky Staiger, Juan Carlos Suarez Serrato, Daniel Xu, and participants of the 2021 conference of the American Society of Health Economists (ASHEcon), the 2021 National Bureau of Economic Research Public Economics and Health Economics Program Meetings, the Harvard–Massachusetts Institute of Technology–Boston University Health Economics Seminar, and the University of Chicago Health Economics Workshop for their helpful feedback. We also thank Paul Kaufman for very helpful discussions about this industry. Evan Henley, Qi Xuan Khoo, Yicheng Liu, Mingrui Ma, Heather Wong, Baiden Wright, and Peck Yang provided excellent research assistance. The data reported here have been supplied by the United States Renal Data System (USRDS). The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the US government. Data were also provided by the Centers for Medicare and Medicaid Services. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health (award nos. P30AG012810 and T32-AG000186). This manuscript was edited by John Asker.

rides to ineligible patients, essentially serving as a very expensive taxi service. From 2003 to 2017, Medicare spent $7.7 billion on 37.5 million nonemergency ambulance rides for dialysis patients.

While the billions of dollars at stake make a study of fraudulent ambulance rides worthwhile on its own, this particular form of fraud represents a larger class of illicit activity in which providers violate Medicare's reimbursement policies by seeking payments for treatments and services without first establishing a medical need for them. A lack of medical necessity has been a key factor in cases as varied as inpatient hospitalizations, physician-administered drugs, nursing homes, and durable medical equipment, amounting to a sizable and preventable waste of Medicare's scarce resources.

The US government uses an array of policies and mechanisms to prevent health-care fraud. One prominent approach, commonly referred to as pay-and-chase litigation, pursues criminal and civil enforcement through the court system, with criminal convictions resulting in jail time and civil judgments imposing heavy penalties on those found guilty of fraud. In contrast to the large expenses incurred by the Federal Bureau of Investigation (FBI) and Department of Justice (DOJ) to investigate and litigate fraud after the fact, the Centers for Medicare and Medicaid Services (CMS) can impose ex ante administrative regulations that restrict reimbursements from being paid out in the first place, such as prior authorization that requires a provider to submit proof of a patient's legitimate medical needs before rendering a service and receiving payment for it.

For our empirical analysis, we use a novel data set of all criminal and civil lawsuits filed against providers accused of ambulance fraud in Medicare's dialysis program over the past two decades combined with Medicare claims data and the staggered rollout of prior authorization across the country to identify the effects of both litigation and regulation on the use of nonemergency ambulance rides, the firms that provide them, patients' access to care, and their resulting health outcomes. We find that adding prior authorization was much more effective at reducing wasteful spending than exclusively pursuing lawsuits against fraudulent providers. Adopting prior authorization caused an immediate and persistent drop in nonemergency ambulance rides of 68%, a substantially larger effect than either criminal or civil litigation had on its own. When weighed against the associated costs of prior authorization and litigation, our results suggest that this type of regulation is an efficient way to reduce unnecessary Medicare expenses.

In addition to causing a large drop in the number of nonemergency ambulance rides to dialysis facilities, prior authorization also led to substantial changes in the market for ambulance services. We find that the number of ambulance companies fell sharply in the markets subject to

prior authorization and that those that remained became more specialized in providing only nonemergency dialysis rides, underscoring important mechanisms through which preemptive regulations can reduce fraud. In line with its limited impact on ridership, however, we find that litigation had a limited effect on the firms not directly prosecuted in the market.

To determine whether the decline in ridership constitutes a reduction in wasteful spending rather than a cut to essential services, we also consider the extent to which prior authorization may have impeded patients' access to care. In this case, the sharp drop in ambulance rides following prior authorization could have made some patients more likely to miss dialysis sessions, increasing their risk of developing serious complications and diminishing their quality of life. Despite this possibility, we find no evidence that the regulatory change disrupted patients' care or led to worse health outcomes, suggesting that prior authorization resulted in a better use of Medicare's resources. We estimate that the federal government would have saved $4.8 billion had it started requiring prior authorization in 2003, when our data begin, rather than waiting until 2014 to pilot the program, and would have done so without any negative health consequences for patients.

We conclude our paper by connecting our empirical results to prominent theories of enforcement and regulation to explain why prior authorization was effective at reducing ambulance fraud while litigation alone was not. A large empirical literature has found that various types of enforcement can effectively reduce criminal behavior and can do so even without an amendment to the underlying law or deliberate change in the probability of enforcement (e.g., Boning et al. 2020). Because the assignment of ex post liability through litigation can deter fraud without incurring up-front monitoring or enforcement costs, the most basic model comparing regulation and litigation would show that regulation is inefficient (Becker 1968; Stigler 1970; Shavell 1984). For ambulance fraud, however, we identify two primary factors undermining the effectiveness of litigation: the limited liability of those committing fraud and their low probability of being detected for it. By directly curtailing the potential gains from providing illegitimate rides, prior authorization does not face the same limitations as litigation and therefore effectively preempts the fraudulent behavior.

We incorporate these insights into a stylized model that connects our findings to past research on limited liability and the likelihood of enforcement, such as Shavell (1984) and Polinsky and Shavell (2000), and extend this literature to a setting where the crime is financial fraud against the government perpetrated by a host of unscrupulous providers. Our work also relates to Glaeser and Shleifer (2003) and Behrer et al. (2021), who consider the trade-offs between regulation and litigation, though the

idea that regulation may be a necessary complement to court enforcement was first considered at least a century ago (Wilson 1913). Connecting these distinct literatures, we believe ours is the first large-scale empirical analysis to study the relative effectiveness of combating fraud by enforcing existing laws or implementing new regulations.[1]

Our findings also add to the literature on fraud and overbilling in Medicare. The seminal works of Silverman and Skinner (2004) and Dafny (2005) lay out the incentives for hospitals to upcode inpatient care to receive larger reimbursements, while Esson (2021) finds that Medicare's rules for establishing medical necessity also lead to upcoding in emergency ambulance services. Others have developed ways to detect suspicious behavior in claims data, such as Fang and Gong (2017), who estimate the time intensity of outpatient procedures to identify providers who bill for an unrealistically large number of hours.[2] Also related is the work of Sanghavi et al. (2021), who link emergency ambulance rides to hospital claims to identify "ghost rides"—rides that do not appear to be substantiated by a hospital visit—among all Medicare beneficiaries, estimating that they make up nearly 2% of rides nationwide. In addition, O'Malley, Bubolz, and Skinner (2021) find that home health-care fraud diffuses faster in cities where firms have more patients in common, while Leder-Luis (forthcoming) studies the economics of civil antifraud health-care litigation conducted against large institutional providers. These studies have largely focused on the incentives to commit fraud and the ways to detect it, which we extend by considering the mechanisms available to combat this type of illicit behavior and the consequences for patients' health.

Our finding that prior authorization reduced spending without harming patients' outcomes relates to the recent debate surrounding administrative burdens in health care (Sahni, Carrus, and Cutler 2021; Brot-Goldberg et al. 2022). In contrast to past work showing that these frictions can limit enrollment (Shepard and Wagner 2025), prompt physicians to stop accepting patients (Dunn et al. 2023), and impose large billing costs on providers without reducing expenditures (League 2023), the modest cost of requiring an ambulance company to obtain prior approval from a physician before transporting a patient to dialysis seems well justified given its success at reducing unnecessary rides and the billions of dollars previously spent on them. In contemporaneous work, a federally funded evaluation study by Contreary, Asher, and Coopersmith

[1] There are several case studies of regulation and litigation in other domains that provide suggestive evidence in favor of one over the other. See, e.g., Harrington, Stockton, and Hooper (2014) or the studies in Kessler (2011). Our paper advances this literature by using modern econometric techniques to identify and quantify the causal impact of each approach in a single, large-scale empirical setting.

[2] Also of note is the related discussion in Matsumoto (2020) and Fang and Gong (2020).

(2022) corroborates our finding that prior authorization reduces Medicare expenditures on nonemergency ambulance rides, though they do not study the corresponding role of litigation, do not consider whether the regulation effectively screens patients who should not be riding in ambulances, and do not investigate its effect on the market for ambulance companies.[3]

Our results suggest that prepayment regulations can be used to curb waste in other federal spending programs where pay-and-chase is the norm, such as the recent wave of fraud in COVID-19 relief aid and the unsuccessful attempts to recover stolen funds (Ackerman and Omeokwe 2022). To this point, the inspector general of the Small Business Administration (SBA) reported that "SBA's lack of adequate front-end controls to determine eligibility contributed to the distribution" of fraudulent loans, making a case for regulations like the type of prior authorization that we study in this paper (SBA 2021). Similar examples of public expenditure fraud abound. The Government Accountability Office (GAO), for instance, estimated that as much as $1.4 billion of Hurricane Katrina relief funds went to improper or fraudulent payments, citing inadequate claim verification as a primary reason (US GAO 2006). Fraud, waste, and abuse in the Iraq reconstruction efforts were estimated at more than $8 billion, with litigation yielding less than $200 million in reclaimed funds (Bowen 2013).

Finally, our paper contributes to a specific literature that has scrutinized the dialysis industry for a host of improper practices. As one example, Eliason et al. (2020) show that independent dialysis facilities acquired by large chains engage in behavior consistent with wasteful drug dumping and increase patients' doses of highly reimbursed drugs, practices found to be detrimental to patients' health. The approach of Fang and Gong (2017) that uses the number of hours worked by a physician to detect overbilling also shows that nephrology is one of the highest categories with claims flagged as infeasible. This literature reflects the pervasive issue of overbilling in dialysis, although not all of it rises to the level of criminal fraud.

Our paper proceeds as follows. Section II discusses the institutional details of dialysis and antifraud enforcement. Section III describes the data and highlights notable descriptive statistics. Section IV outlines our empirical framework. Section V presents our empirical results, including the effects of prior authorization and litigation on nonemergency ambulance rides, the firms that provide them, patients' health outcomes, and

[3] Contreary, Asher, and Coopersmith (2022) was submitted in February 2022 and cites our November 2021 National Bureau of Economic Research working paper version of this article.

the subsequent characteristics of riders. Section VI develops a stylized model to orient our empirical findings within the theoretical literature studying the effectiveness of regulation and litigation. Section VII concludes with our arguments for why regulatory actions are a cost-effective way to prevent health-care fraud.

II. Background

Medicare's End-Stage Renal Disease (ESRD) program covers patients needing dialysis, a procedure that cleans the blood of those without well-functioning kidneys. Dialysis patients typically visit one of the nation's more than 7,000 dialysis facilities three times per week to receive treatments that last 3–4 hours each session. Many patients arrange for transportation to dialysis on their own, either in a personal vehicle or on public transit, but those with severe medical conditions require an ambulance. Medicare pays for transportation to and from dialysis only when an ambulance is medically necessary, meaning that the patient has no other safe way to travel due to their medical condition.[4]

Ambulance companies must satisfy several requirements to receive Medicare reimbursements for providing rides to dialysis facilities. Federal regulations stipulate that ambulances must be staffed by at least two people, with at least one certified as an emergency medical technician, and that the vehicles must be specifically designed as ambulances.[5] In addition, providers need a National Provider Identifier (NPI), and dialysis patients must be bedridden or need lifesaving procedures in transit for the ride to qualify as medically necessary.[6]

Medicare pays for ambulance rides through Part B, making patients responsible for a 20% copayment on top of their annual deductible. The payment rates consist of a base fee, which depends on the level of life

[4] The Medicare Benefit Policy Manual specifies that "in any case in which some means of transportation other than an ambulance could be used without endangering the individual's health, whether or not such other transportation is actually available, no payment may be made for ambulance services." Submitting claims for care that fails to meet the medical necessity standard constitutes health-care fraud.

[5] States may also impose their own regulations, such as the certificate of need laws currently in place in Arizona, Hawaii, Iowa, Kentucky, New Jersey, and New York. All states also license various levels of emergency medical service occupations and have different requirements for these licenses.

[6] The Code of Federal Regulations, Title 42, Chapter IV, Part 410.40, stipulates, "Nonemergency transportation by ambulance is appropriate if either: the beneficiary is bed-confined, and it is documented that the beneficiary's condition is such that other methods of transportation are contraindicated; or if his or her medical condition, regardless of bed confinement, is such that transportation by ambulance is medically required. . . . For a beneficiary to be considered bed-confined, the following criteria must be met: (i) The beneficiary is unable to get up from bed without assistance. (ii) The beneficiary is unable to ambulate. (iii) The beneficiary is unable to sit in a chair or wheelchair."

support (e.g., whether the ride was an emergency or, in rare cases, required air transportation) and a per-mile fee, for which ambulances receive a bonus if the pickup is in a rural location. For the nonemergency ground transportation that we focus on in this paper, the current base and mileage rates are $272.44 and $8.76, respectively, up from $209.65 and $6.74 in 2010, with rates adjusted by location.

Fraud has become a major concern for all of Medicare's ambulance reimbursements, not just among dialysis patients. The Department of Health and Human Services Office of Inspector General (HHS-OIG) has published several reports about Medicare's ambulance benefit, including "Medicare Payment for Ambulance Transport" (HHS-OIG 2006), which found that 20% of nonemergency transports were improper in that they did not meet Medicare's coverage requirements.

The issue is particularly acute in dialysis, however, where for many years ambulance companies transported patients who did not meet Medicare's criteria for receiving medically necessary rides. The large reimbursements paid by Medicare, coupled with patients' regularly scheduled and recurring visits to facilities, create a strong financial incentive for unscrupulous providers to engage in fraud, especially if they transport nonemergency patients who do not require costly medical attention during the ride. From 2007 to 2011, the volume of transports to and from dialysis facilities increased by more than twice the rate of all other ambulance transports. In 2011, ambulance rides to and from dialysis facilities accounted for nearly $700 million in Medicare spending, or approximately 13% of Medicare's total expenditures on ambulance services (CMS 2020b). Reflecting this growth, figure 1 shows the number of rides in our data more than tripling from 2003 to 2014, a period when the number of ESRD patients increased by only 54%.

The US government has used several different approaches to prevent unnecessary ambulance rides for dialysis patients. Those who commit Medicare fraud can run afoul of criminal statutes, including the health-care fraud statute (18 U.S.C. §1347) and the antikickback statute (42 U.S.C. §1320a-7b(b)), with the crimes investigated by the FBI and prosecuted by DOJ district offices nationwide. The United States compounds its enforcement with laws against conspiracy, racketeering, organized crime, and lying to investigators. Employing this pay-and-chase approach, over the past 25 years the DOJ has pursued 43 criminal lawsuits against ambulance company operators for providing fraudulent rides to dialysis patients, alleging illegal behavior like paying kickbacks to patients to induce them to ride, giving referral bonuses to patients who recruited others to participate in the scheme, and concealing or manipulating documentation to justify the ongoing use of ambulances.

In addition to criminal statutes, federal health-care fraud violates the False Claims Act, a civil statute that imposes monetary penalties of up to

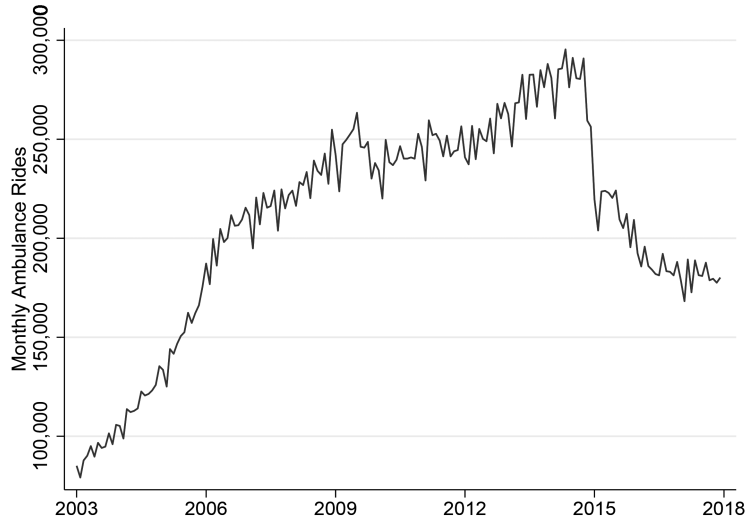

FIG. 1.—Nonemergency basic life-support dialysis rides over time. The sample includes nonemergency basic life-support ambulance rides from a dialysis facility to a place of residence for ESRD patients from 2003 to 2017.

triple damages on firms that overbill federal health-care programs. The False Claims Act contains a qui tam whistleblower provision where individuals with knowledge and evidence of fraud can file their own lawsuits against those who submit fraudulent claims on behalf of the US government in exchange for 15%–30% of the recovered funds, while the DOJ can also initiate civil lawsuits on its own. We identify 26 civil lawsuits, from as early as 1996, alleging the unnecessary transport of dialysis patients by ambulance companies.

Medicare administrators also attempt to stop overbilling and fraud by enacting new regulations. In the case of medically unnecessary ambulance rides, Medicare began implementing prior authorization for ambulance claims in 2014, stipulating that providers will only receive payment for repetitive, nonemergency rides to dialysis facilities if they have already submitted documentation of a patient's medical necessity, rather than allowing providers to submit claims for payment first and then responding to any subsequent requests to verify a patient's eligibility. Medicare began rolling out the new requirement in 2014 in three states that had particularly high rates of nonemergency ambulance claims—New Jersey, South Carolina, and Pennsylvania—and then extended it in 2016 to nearby Delaware, District of Columbia, Maryland, North Carolina, Virginia, and West Virginia. Plans to expand prior authorization nationwide were postponed in 2020 due to COVID-19 but eventually completed in August 2022, with policymakers still debating the merits of the regulation (Lotven 2022).

III. Data and Descriptive Statistics

We use the 100% sample of claims data compiled by the United States Renal Data System (USRDS) for the entire universe of patients diagnosed with ESRD and enrolled in Medicare between 2003 and 2017.[7] The patient-level data allow us to observe demographics (e.g., sex, race, body mass index, cause of ESRD, payer, comorbidities, zip code, and a facility identifier) and complete ESRD treatment histories, while the facility-level data have information on location and ownership. Our data also allow us to observe each ambulance ride to and from a dialysis facility billed to Medicare, which amounts to more than 37.5 million nonemergency rides and more than $7.7 billion in spending. For firms that provide nonemergency ambulance rides, we have additional data on their other claims for Medicare ESRD beneficiaries, such as emergency hospital transports. In the past 6 years of our data alone, we observe 3,081 firms providing nonemergency rides to dialysis patients. Because the USRDS data only began recording firm identifiers in 2012, we supplement these data with a 20% sample of claims for all Medicare beneficiaries between 2007 and 2019.[8]

Table 1 provides summary statistics for patient characteristics, ridership, and health outcomes for those who receive any nonemergency ride to a dialysis facility, split across months with and without rides, as well as summary statistics for dialysis patients who never receive such a ride. Riders are older, more likely to be women, more likely to be Black, and more likely to have diabetes. Patients who use ambulances for nonemergency transportation to dialysis facilities take 10 round-trip rides each month, on average, amounting to 20 claims total, with a lifetime average of 660 claims. Because dialysis patients receive approximately 12 treatments per month, these averages imply that patients who take an ambulance to and from their facility do so for nearly 9 out of 10 sessions.

We supplement these data with information on criminal and civil enforcement against fraud. Using publicly available press releases from the DOJ, corroborated for completeness by internet searches, we identify 69 lawsuits across 26 federal judicial districts against dozens of ambulance companies and individuals for unnecessary ambulance rides related to dialysis. For each of these lawsuits, we collect court records from the Public Access to Court Electronic Records (PACER) system, which include

[7] USRDS combines data from a variety of sources, including Medicare claims, annual facility surveys, and dialysis treatment histories, to create the most comprehensive data set for studying the US dialysis industry. For a more thorough description of USRDS, please see the *Researcher's Guide to the USRDS System* (USRDS 2020).

[8] Because the USRDS data are a 100% sample of claims, we use this as our primary data source, relying on the 20% sample only when assessing firm-level outcomes related to criminal and civil litigation, which often occurred before 2012. Unless otherwise noted, all calculations, tables, and figures rely on the USRDS data.

TABLE 1
Summary Statistics of Patient-Month Data

	Patient Rider Status			
		Rider		
	Never-Rider	Nonriding Month	Riding Month	Overall
Patient characteristics:				
Age (years)	61.97	67.20	69.27	62.99
Months with ESRD	55.99	60.38	54.05	56.49
Black	.377	.418	.451	.386
Male	.561	.496	.457	.548
Diabetic	.524	.620	.661	.543
Drug user	.014	.011	.008	.013
Smoker	.065	.055	.045	.063
Drinker	.013	.013	.011	.013
Uninsured at incidence	.129	.089	.061	.120
Employed at incidence	.181	.099	.066	.165
Ridership:				
Nonemergency dialysis rides	.00	.00	19.54	.87
Emergency rides	.100	.179	.408	.125
Total lifetime rides	.0	132.8	660.3	47.1
Continuing to ride next month	.	.	.838	.838
Facility characteristics				
Facility age (years)	16.94	16.48	16.30	16.85
Freestanding facility	.956	.966	.972	.958
Chain affiliation:				
DaVita	.345	.333	.325	.343
Fresenius	.364	.372	.377	.366
Other	.135	.148	.153	.137
Independent	.156	.146	.146	.154
Health outcomes:				
Dialysis sessions	12.18	12.05	11.29	12.13
All-cause hospitalizations	.110	.152	.250	.122
Fluid hospitalizations	.011	.016	.020	.012
Mortality	.009	.007	.034	.010
Patient-months	15,611,284	2,533,118	846,573	18,990,975

Note.—Data are from 2011 to 2017. Patient characteristics except age and dialysis tenure are at incidence of ESRD. All ridership variables other than emergency rides are based on nonemergency basic life-support rides between a dialysis facility and a patient's home. The probability of continuing to ride is the conditional probability of riding in the next month given the patient rides in the focal month. Fluid hospitalizations are those for which the primary diagnosis indicates excess fluids, an indication of insufficient dialysis.

specific fraud allegations and data on the lawsuit's timing and location of enforcement.[9]

As discussed in section II, Medicare's regulation requiring prior authorization stipulates that ambulance companies must obtain approval for

[9] We use the court filing or complaint date as the treatment date. Civil lawsuits are often filed under seal, meaning it is unlikely that the lawsuit's existence was known prior to the filing date. Criminal lawsuits may involve investigations before the lawsuit is filed and there

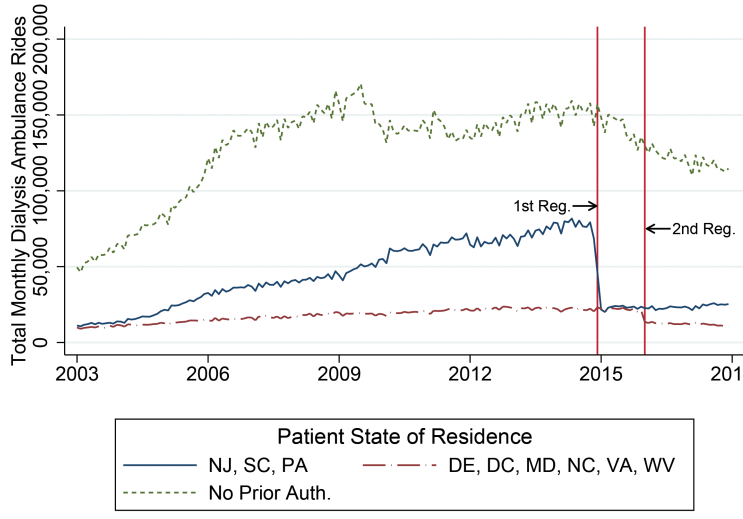

FIG. 2.—Rides by prior authorization regulation. For each of the three lines, the vertical axis measures total rides per month in the represented states. Sample includes nonemergency basic life-support ambulance rides from a dialysis facility to a place of residence for dialysis patients from 2003 to 2017. State determined by the transported patient's residence. The first vertical line marks the start of prior authorization in New Jersey, South Carolina, and Pennsylvania, and the second marks that in Delaware, Washington, DC, Maryland, North Carolina, Virginia, and West Virginia.

each patient receiving repetitive, nonemergency ambulance transports before they provide the service, with the approval renewed periodically.[10] This policy was piloted on December 15, 2014, in New Jersey, Pennsylvania, and South Carolina and then expanded on January 1, 2016, to Delaware, District of Columbia, Maryland, North Carolina, Virginia, and West Virginia. Figure 2 shows preliminary evidence of the regulation's effectiveness: rides for patients in Pennsylvania, New Jersey, and South Carolina fell sharply after Medicare first imposed prior authorization, with states included in the second wave experiencing a similar decline immediately upon the policy's expansion. Figures A2 and A3 (figs. A1–A23 are

is some chance that firms become aware of them prior to the filing date. However, in both cases, this date likely represents the best-possible case for this analysis. It is also the standard date used in the literature on health-care fraud and beyond (Agan, Freedman, and Owens 2021; Agan, Doleac, and Harvey 2023; Gruber et al. 2025; Leder-Luis, forthcoming). Furthermore, our event study design allows us to demonstrate the robustness of this decision. Appendix A provides more detail about these data, including specifics about the timing of the litigation activity.

[10] Medicare considers "three or more round trips during a 10-day period, or at least one round trip per week for at least three weeks" to be repetitive transports (CMS 2020b). Prior authorization is required for the fourth ride in a 30-day period.

available in the online appendixes) show the raw trends for nonemergency dialysis ambulance rides by district surrounding the timing of their lawsuits.

IV. Empirical Strategy

We use the staggered rollout of prior authorization and the differential timing of criminal and civil enforcement across US federal judicial districts to identify the causal effects of these respective approaches for reducing unnecessary rides and their impact on patients.[11] For our estimates, we present results using both traditional two-way fixed effects (TWFE) methods in the main text and several alternative estimators in appendix B (apps. A–K are available online). For the traditional TWFE results, we estimate

$$Y_{dt} = \sum_{e=-K}^{-2} \beta_e T_{dt}(e) + \sum_{e=0}^{L} \beta_e T_{dt}(e) + \alpha_d + \alpha_t + \Gamma X_{dt} + \epsilon_{dt} \quad (1)$$

for district d in month t, where Y_{dt} is the outcome of interest (e.g., payments or rides, measured in both levels and logs), $T_{dt}(e)$ is an indicator for an observation falling e months from the treatment date with base period $e = -1$, α_d and α_t are district and month fixed effects, and X_{dt} is a matrix of indicators for having already been subject to a different type of enforcement or prior authorization. Because districts are geographic subsets of states, district fixed effects account for state fixed effects.

To avoid the compositional issues that have been noted by, for example, Callaway and Sant'Anna (2021), we set $K = 24$ and $L = 23$, defining $T_{dt}(e)$ only for units in the sample for the entire 48-month period around the treatment date and only for observations in that window. For untreated units, we set $T_{dt}(e) = 0$ for all e. We also use alternative estimators that directly address compositional issues in appendix B.

To aggregate these results into a single parameter, we estimate

$$Y_{dt} = \sum_{e=-K}^{-2} \beta_e T_{dt}(e) + \beta \sum_{e=0}^{L} T_{dt}(e) + \alpha_d + \alpha_t + \Gamma X_{dt} + \epsilon_{dt}. \quad (2)$$

This is similar to the more traditional pre-post estimator, but rather than comparing the entire preperiod to the entire postperiod, the entire postperiod is compared only with the period immediately before treatment (i.e., $e = -1$). This estimator explicitly captures the average treatment

[11] There are 94 US federal judicial districts, each of which is wholly contained within a state; these are the regions at which the Department of Justice and the US federal courts operate, each with its own US Attorney and Department of Justice office. We provide a map of these districts in app. A.

effect on the treated over the first L months of treatment rather than the varying lengths of time captured by a pre-post indicator, which potentially could be quite different. By setting $K = 24$ and $L = 23$, we capture the effect of treatment in the 2 years following treatment.

For results estimated at the patient level, our estimating equations are

$$Y_{idt} = \sum_{e=-K}^{-2} \beta_e T_{dt}(e) + \sum_{e=0}^{L} \beta_e T_{dt}(e) + \alpha_d + \alpha_t + \Gamma X_{idt} + \epsilon_{idt} \quad (3)$$

and

$$Y_{idt} = \sum_{e=-K}^{-2} \beta_e T_{dt}(e) + \beta \sum_{e=0}^{L} T_{dt}(e) + \alpha_d + \alpha_t + \Gamma X_{idt} + \epsilon_{idt} \quad (4)$$

for individual i with observable patient and dialysis facility characteristics X_{idt}. Here, we set $K = 12$ and $L = 11$ to capture the effect over the first year.

For further justification of our research design, we provide a balance table comparing control states to prior authorization states by each wave of the regulation's rollout in table A8 (tables A1–A33 are in the online appendixes). Although some small differences exist, the health outcomes are similar in terms of hospitalization and mortality rates as well as the rate of emergency ambulance rides. The second-wave states are also similar to the control states in terms of nonemergency ridership, although the first-wave states did have higher ridership overall. Similarly, table A9 shows that observable characteristics are balanced across districts subject to prior authorization and litigation, supporting our choice to compare the effects of each intervention.

Finally, we perform several robustness checks using alternative difference-in-differences estimators suggested by recent developments in the literature on using TWFE estimators with staggered treatments or heterogeneous treatment effects (Borusyak, Jaravel, and Spiess 2017; Cengiz et al. 2019; Callaway and Sant'Anna 2021). Because the traditional TWFE approach relies solely on within-group variation in the treatment variable to eliminate possible unobserved confounders related to districts or time trends, staggered treatment timing may result in inappropriate comparisons, such as including already treated districts as controls, and potentially bias our estimates. In light of this concern, we show in appendix B that none of the alternative difference-in-differences estimators affect our results, while in appendix D we show our results are robust to using alternative control groups (e.g., using only not-yet-treated districts or only bordering districts as controls).

Our empirical strategy allows us to identify three important, policy-relevant parameters: the average treatment effects of (i) adding prior authorization, (ii) pursuing criminal litigation, and (iii) pursuing civil litigation. Because the possibility of litigating fraudulent ambulance

companies always exists throughout our sample period, we cannot consider the impact of imposing a new litigation regime. Instead, our empirical design compares a policymaker's two primary options when current enforcement mechanisms do not deter fraud effectively: pursuing litigation to enforce existing laws or implementing new regulations that make fraud less lucrative.

V. Empirical Results

A. *Payments and Rides*

We first consider the effect of prior authorization on rides and spending. Table 2 provides estimates of the policy's effect on the number of nonemergency ambulance rides between a dialysis facility and a patient's home, as well as Medicare payments for such rides, in all treated districts in the 2 years following treatment, as represented by β in equation (2). Outcomes are measured both in levels and by adding one and taking the natural log.

We find that prior authorization reduces payments for nonemergency ambulance rides by 1.129 log points, or 67.7%.[12] Figure 3 shows the dynamic difference-in-differences results, or estimates of β_e for $e \in [-24, 23]/\{-1\}$ in equation (1), with log-transformed total payments as the dependent variable. We find that the effect of prior authorization was large, immediate, and persistent.[13]

In contrast to prior authorization's requirement that providers seek approval before receiving payment, both criminal and civil litigation attempt to identify and prosecute illicit behavior after it has already taken place, a pay-and-chase approach aimed at both deterring fraud and punishing those who commit it. Under this system, the threat of litigation is always present, although successful litigation may act as an even stronger deterrent by changing the incentives for those who might commit similar fraudulent acts (Leder-Luis, forthcoming). For that reason, we estimate the impact of realized litigation in the local district rather than the most general deterrence that would arise from having a law already in place or enacting a new one.

[12] The second wave of the prior authorization rollout occurred 2 years before the end of our data, meaning that both treatment waves are included in this parameter. To address the possibility that this masks meaningful differences in the effect across the two waves, we also estimate separate treatment effects for each wave and show the results in app. E. We find a reduction in payments of 1.21 log points in the first-wave states and 1.07 log points in the second-wave states. The difference between these two estimates is not statistically significant.

[13] In app. F, we perform a similar analysis at the firm-month and patient-month levels, finding that the large effect of prior authorization is robust. We also show that our results are robust to using the inverse hyperbolic sine transformation or a Poisson specification. Finally, we also consider a falsification test that shows prior authorization had no impact on the number of emergency rides.

TABLE 2
EFFECT OF PRIOR AUTHORIZATION ON AMBULANCE RIDES AND SPENDING

	Total Ride Payments (Log) (1)	Total Ride Payments (2)	Total Rides (Log) (3)	Total Rides (4)
Prior authorization	−1.129**	−738674.2^{+}	−.913***	−3,714.6^{+}
	(.350)	(405,698.1)	(.176)	(2,039.7)
Year-month fixed effects	1	1	1	1
District fixed effects	1	1	1	1
Dependent variable mean	9.934	415,286.7	5.357	2,005.3
Observations	7,272	7,272	7,272	7,272

NOTE.—Estimates of β from eq. (2). All rides are nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data. Dependent variable in cols. 1 and 3 are transformed by adding 1 and taking the natural log. These data include rides from 2011 to 2017. An observation is a district-month. Standard errors are clustered at the district level.

$^{+}$ Significance at the 10% level.

** Significance at the 1% level.

*** Significance at the 0.1% level.

To study the impact of litigation on ambulance fraud, we use the same approach as above for civil and criminal enforcement actions.[14] Table 3 provides estimates of β from equation (2), where the treatment date is determined by the start of each type of enforcement in the district.[15]

We find that civil enforcement does not have a statistically significant effect on rides or total payments, whereas criminal enforcement reduces monthly payments by 19% and rides by 24% in the subsequent two years.[16] Figure 4 shows the dynamic effects of the first indictment of each type. Although we find no decrease in payments following civil enforcement, our results suggest that criminal enforcement gradually reduces payments over time.

[14] This methodology relies on districts that are not subject to enforcement serving as a reliable comparison group for those that are. In particular, if there are national or regional spillovers in the effect of indictments beyond the districts in which they occur, our estimates would be biased. In app. G, we show that the effects of enforcement are highly localized, with no negative impacts on rides or payments in neighboring districts.

[15] Because Illinois North, Massachusetts, Arkansas East, North Carolina East, and California Central had civil actions before or within the first year of our sample period and the first civil actions in Georgia South and Virginia East occurred too late in our data, we exclude these districts from our analysis of civil enforcement. Similarly, Arkansas East, California Central, and North Carolina East are excluded from our analysis of criminal enforcement because the associated actions occurred too early in our data, while Kentucky East is excluded because its enforcement actions occurred too late.

[16] In app. D, we show that these results are robust to alternative functional form assumptions and control groups. In app. E, we investigate a number of potential dimensions of heterogeneity in this effect that may indicate endogenous enforcement, including heterogeneity by enforcement date, prelitigation ridership, and the number of cases pursued in the DOJ district. We find little evidence of such heterogeneity. In app. D, we also estimate a single specification that includes both litigation and regulation, finding effects similar to those above.

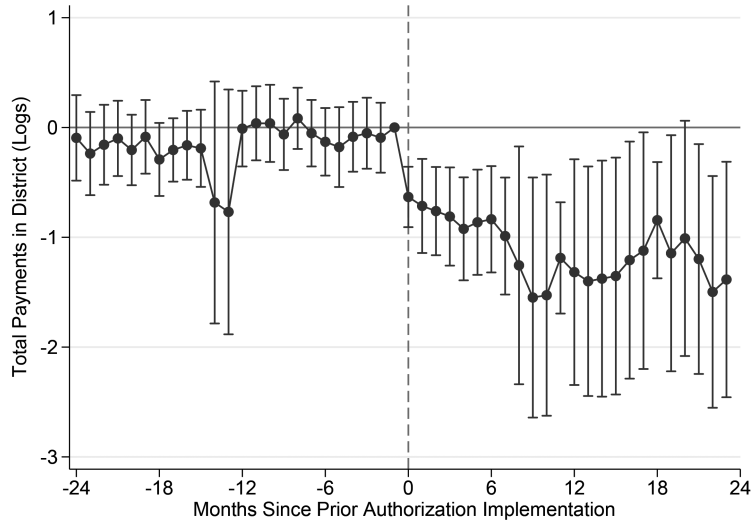

FIG. 3.—Effect of prior authorization on ambulance spending: estimates of β_e for $e \in [-24, 23]/\{-1\}$ from equation (1). Dependent variable is total payments for nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data transformed by adding 1 and taking the natural log. These data include rides from 2011 to 2017. An observation is a district-month. Standard errors are clustered at the district level. Error bars represent pointwise 95% confidence intervals.

The relative magnitudes of these enforcement approaches are highlighted in figure 5, which presents the estimates from figures 3 and 4 in a single panel to illustrate the stark difference between the effect of prior authorization and the effects of both criminal and civil litigation. Even 2 years after the enforcement action, our results suggest that criminal enforcement has only 20%–25% of the effect of prior authorization and that civil enforcement continues to have no impact whatsoever.[17] In appendix H, we present analogous figures for other outcomes to demonstrate the notable differences across each type of intervention. In all cases, the impact of prior authorization is qualitatively much larger than litigation and the differences are statistically significant.

B. *Patient Health*

Although prior authorization reduced the number of ambulance rides taken by dialysis patients, the additional administrative burden may have

[17] A natural question is whether litigation would have a larger effect if multiple cases were brought in a district. While few districts have multiple cases, we present evidence in table A20 that there are not large deterrence effects from cases subsequent to the first in a district.

TABLE 3
EFFECT OF LITIGATION ON AMBULANCE SPENDING AND RIDES

	Civil		Criminal	
	Total Ride Payments (Log) (1)	Total Rides (Log) (2)	Total Ride Payments (Log) (3)	Total Rides (Log) (4)
Enforcement	−.0424	.0257	−.211+	−.280**
	(.110)	(.0663)	(.106)	(.0994)
Year-month fixed effects	1	1	1	1
District fixed effects	1	1	1	1
Dependent variable mean	9.221	4.835	9.354	4.928
Observations	14,160	14,160	14,436	14,436

NOTE.—Estimates of β from eq. (2). All rides are nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data. Dependent variables are transformed by adding 1 and taking the natural log. These data include rides from 2003 to 2017. An observation is a district-month. The treatment date is the earliest enforcement action of the relevant type in the district. Standard errors are clustered at the district level.

+ Significance at the 10% level.

** Significance at the 1% level.

resulted in some patients forgoing treatment if they could not find another safe way to reach their facilities. If these missed sessions resulted in adverse events such as hospitalization or death, Medicare's savings from fewer ambulance reimbursements could have been offset by higher costs in other parts of the ESRD program, to say nothing of the lower quality of life for the affected patients.

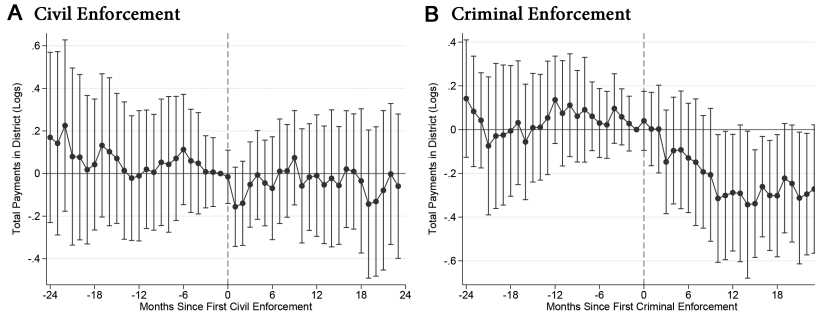

FIG. 4.—Impact of litigation on ambulance payments: estimates of β_e for $e \in [-24, 23]/\{-1\}$ from equation (1). Dependent variable is total payments for nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data transformed by adding 1 and taking the natural log. These data include rides from 2003 to 2017. An observation is a district-month. The treatment date is the earliest enforcement action of the relevant type in the district. Standard errors are clustered at the district level. Error bars represent pointwise 95% confidence intervals.

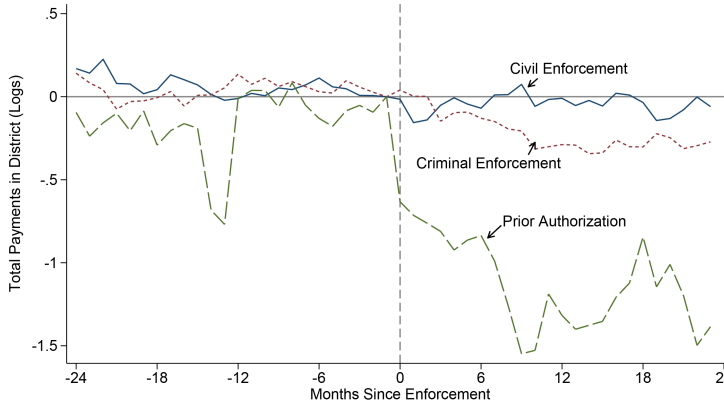

FIG. 5.—Effect of prior authorization and criminal and civil litigation on ambulance payments: estimates of $e \in [-24, 23]/\{-1\}$. Dependent variable is total payments for nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data transformed by adding 1 and taking the natural log. These data include rides from 2003 to 2017. An observation is a district-month. The treatment date is the earliest enforcement action of the relevant type in the district.

To assess the impact of prior authorization on health outcomes, we estimate equation (4) at the patient-month level, with measures of patients' health as the outcome variables. We control for a rich set of patient and facility characteristics and include facility fixed effects while clustering standard errors at the district level.

Table 4 presents the effects of prior authorization on patients' adherence to dialysis and on downstream health outcomes such as hospitalizations and mortality. We find no evidence that prior authorization led to either meaningful decreases in dialysis sessions or increases in adverse events, ruling out even a 0.6% decrease in monthly dialysis sessions at the 95% confidence level.

Although we do not find that prior authorization harmed patients' health on average, it could be that some patients were harmed in ways not captured by our point estimates. To rule out this possibility, we restrict our sample to the group of patients most likely to be affected by the policy change: those who relied most heavily on ambulance rides prior to the reform. Specifically, we restrict our sample to patients who took at least 100 nonemergency ambulance rides to dialysis facilities before prior authorization and compare the outcomes of these frequent riders throughout the staggered rollout of prior authorization across districts. Table 5

[17] A natural question is whether litigation would have a larger effect if multiple cases were brought in a district. While few districts have multiple cases, we present evidence in table A20 that there are not large deterrence effects from cases subsequent to the first in a district.

TABLE 4
Effect of Prior Authorization on Adherence and Adverse Events

	Dialysis Sessions (1)	Mortality (2)	All-Cause Hospitalizations (3)	Fluid Hospitalizations (4)
Prior authorization	−.0256	.000372	−.00132	−.000854
	(.0191)	(.000580)	(.00136)	(.000777)
Year-month fixed effects	1	1	1	1
District fixed effects	1	1	1	1
Patient/facility controls	1	1	1	1
Facility fixed effects	1	1	1	1
R^2	.0108	.00493	.0159	.00610
Dependent variable mean	12.12	.00988	.122	.0116
Observations	15,077,158	15,077,158	15,077,158	15,077,158

Note.—Estimates of β from eq. (4) at the patient-month level. Data are from 2011 to 2017. Controls include incident patient characteristics, age, and tenure on dialysis as well as facility fixed effects and facility characteristics including chain ownership status, demographic characteristics of the zip code, and whether the facility is freestanding or hospital based. Fluid hospitalizations are those for which the primary diagnosis indicates fluid overload, often an indication of insufficient dialysis. Standard errors clustered at the district level are given in parentheses.

shows that, even for the most frequent riders, nothing suggests prior authorization resulted in worse health outcomes.

We also find no evidence of meaningful changes in patients' health following criminal and civil litigation, as shown in tables A27–A28. With litigation affecting ridership much less than prior authorization did, it is

TABLE 5
Effect of Prior Authorization on Frequent Riders

	Dialysis Sessions (1)	Mortality (2)	All-Cause Hospitalizations (3)	Fluid Hospitalizations (4)
Prior authorization	−.0226	−.000433	−.00828	−.00137
	(.0312)	(.00167)	(.00517)	(.00176)
Year-month fixed effects	1	1	1	1
District fixed effects	1	1	1	1
Patient/facility controls	1	1	1	1
Facility fixed effects	1	1	1	1
R^2	.0742	.0109	.0281	.0148
Dependent variable mean	11.88	.0115	.179	.0155
Observations	905,331	905,331	905,331	905,331

Note.—Estimates of β from eq. (4) at the patient-month level. Data are from 2011 to 2017. Controls include incident patient characteristics, age, and tenure on dialysis as well as facility fixed effects and facility characteristics including chain ownership status, demographic characteristics of the zip code, and whether the facility is freestanding or hospital based. Fluid hospitalizations are those for which the primary diagnosis indicates fluid overload, often an indication of insufficient dialysis. The sample is limited to patients who took at least 100 nonemergency ambulance rides to dialysis under the non–prior authorization regime. Standard errors clustered at the district level are given in parentheses.

perhaps not surprising that we similarly see no impact on health outcomes for both of these measures as well.

C. *Mechanisms of Prior Authorization*

Not only did prior authorization cause a large drop in the number of nonemergency ambulance rides to dialysis facilities, it also led to substantial changes in the underlying market for ambulance services. As shown in table 6 and the corresponding event study in figure 6, prior authorization resulted in an abrupt reduction in the number of ambulance companies providing nonemergency dialysis rides by 0.286 log points, or 24.9%. We find that, beyond simply reducing the number of ambulance companies, prior authorization also led to greater firm specialization: firms with a higher share of nonemergency rides were more likely to exit following the first wave of prior authorization's rollout, while the number of firms providing only nonemergency dialysis rides increased. The distribution of firms broken down by the share of nonemergency rides that they provide to dialysis patients in figure 7 shows that many of the firms that provide nonemergency ambulance rides to dialysis patients provide very few emergency rides to the same population. After prior authorization, fewer firms provide nonemergency rides to dialysis patients overall, but the effect is most pronounced among firms that provide a moderate share of nonemergency rides. At the same time, the number of firms providing only nonemergency rides to dialysis patients increased by a third, from 93 to 120, indicating that regulation led to specialization among the firms that continued to provide this service.

A within-firm analysis provides further evidence of specialization following prior authorization. Firms that initially provided few nonemergency

TABLE 6
EFFECT OF PRIOR AUTHORIZATION ON NUMBER OF ACTIVE FIRMS

	Active Firms (Log) (1)	Active Firms (2)
Prior authorization	−.286***	−13.96*
	(.0657)	(5.906)
Year-month fixed effects	1	1
District fixed effects	1	1
Dependent variable mean	2.152	17.23
Observations	6,336	6,336

NOTE.—Estimates of β from eq. (2). Dependent variables are the number of firms providing nonemergency basic life-support rides between a dialysis facility and a patient's home in a district-month and the natural logarithm of 1 plus the same. These data include rides from 2012 to 2017. An observation is a district-month. Standard errors are clustered at the district level.

* Significance at the 5% level.

*** Significance at the 0.1% level.

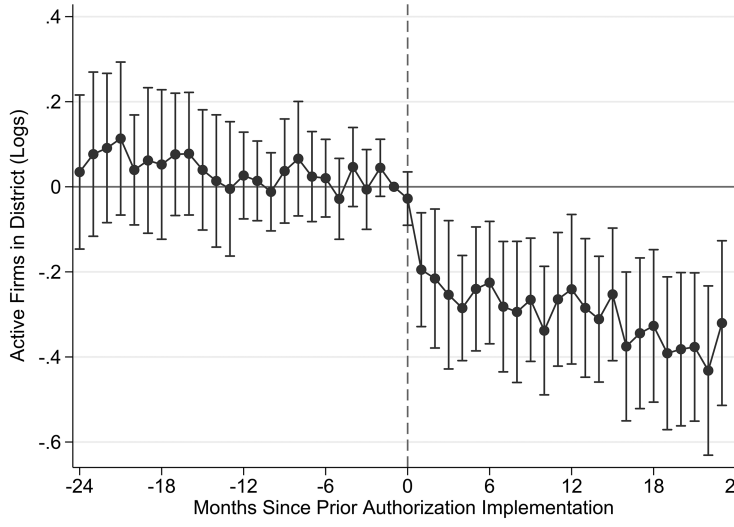

FIG. 6.—Effect of prior authorization on the number of active firms: estimates of β_e for $e \in [-24, 23]/\{-1\}$ from equation (1). Dependent variable is the number of firms providing nonemergency basic life-support rides between a dialysis facility and a patient's home in a district-month transformed by adding 1 and taking the natural log. These data include rides from 2012 to 2017. An observation is a district-month. Standard errors are clustered at the district level. Error bars represent pointwise 95% confidence intervals.

rides were much more likely to stop providing rides altogether after prior authorization: over half of the firms for which nonemergency dialysis rides comprised less than 20% of their total rides no longer provide the service at all. At the other extreme, firms more concentrated in nonemergency rides before the regulation were much less likely to exit this market following prior authorization and in some cases began to specialize even more in providing them, as shown in figure A20.

Prior authorization may also reduce fraudulent activity by ensuring that only patients who qualify for rides under Medicare's reimbursement policy end up receiving them. To qualify for a nonemergency ambulance ride, a dialysis patient must be unable to travel safely by any other means, as in the case of a permanently bedridden patient or one who needs a short stint of rides following a hospitalization. In contrast to litigation that only targets fraudulent behavior, the success of prior authorization depends on its ability to deter fraudulent rides while at the same time not deterring legitimate ones. Despite this delicate trade-off, several stylized facts suggest that the regulation achieved its primary aim of reducing unnecessary rides without discouraging those who truly need them.

First, we find that prior authorization led not only to fewer riders overall but also to less persistent and shorter ridership spells among those who

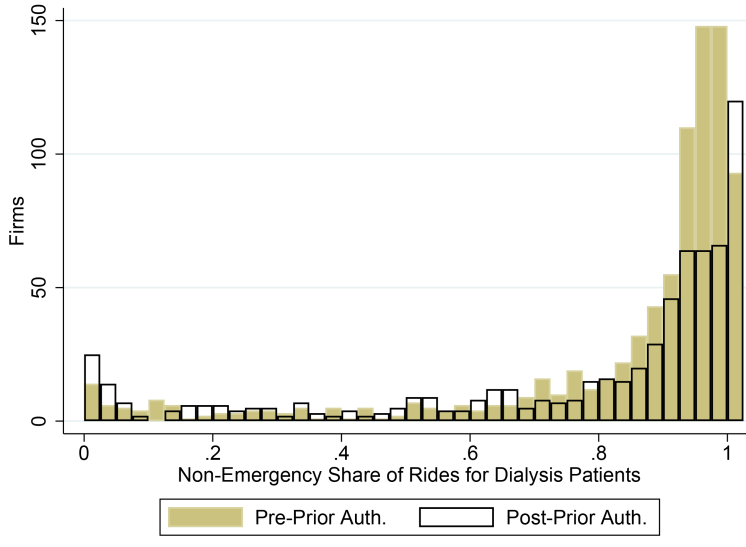

FIG. 7.—Change in distribution of firms by share of nonemergency rides: the distribution of ambulance firms that served dialysis patients in the 3 years before and after prior authorization in states subject to prior authorization in December 2014. A firm's pre–prior authorization nonemergency share is determined by the share of total rides given by the firm in the 36 months before the start of prior authorization in that state that were nonemergency rides between a dialysis treatment facility and a patient's residence. The post–prior authorization share is the same share for the 36 months following the implementation of prior authorization. Firms that gave no nonemergency dialysis rides in the relevant period are excluded.

take ambulances, a result consistent with the benefit being used predominately by acutely ill patients who ride for only a limited time. As shown in column 1 of table 7, which contains estimates of equation (4) with different outcome variables restricted to patients taking an ambulance in the current month, the probability that a current rider continues riding in the following month fell after prior authorization, indicating that ridership became less persistent. Also consistent with this interpretation, the median number of months in which a rider takes a nonemergency ride fell from six to three and the total number of rides taken by each rider decreased substantially in the 2 years after prior authorization compared with the 2 years immediately preceding it, as shown in figure 8.

In addition to reducing the duration of ridership spells, prior authorization also resulted in rides being targeted to patients in poorer health. Columns 2 and 3 of table 7 and panels B and C of figure 9 show that the share of ambulance riders suffering an adverse event in the same month they take a ride increased after prior authorization, suggesting a larger proportion of riders with a legitimate need for an ambulance. Taken as a whole, these results indicate that the patients receiving nonemergency

TABLE 7
EFFECT OF PRIOR AUTHORIZATION ON PATIENT SELECTION

	Rides Next Month (1)	Hospitalizations (2)	Mortality (3)
Prior authorization	−.0633	.0117+	.00711*
	(.0529)	(.00630)	(.00348)
Year-month fixed effects	1	1	1
District fixed effects	1	1	1
Patient/facility controls	1	1	1
Facility fixed effects	1	1	1
R^2	.113	.0422	.0239
Dependent variable mean	.829	.256	.0352
Observations	603,917	603,917	603,917

NOTE.—Estimates of β from eq. (4) at the patient-month level. Data are from 2011 to 2017. Controls include incident patient characteristics, age, and tenure on dialysis as well as facility fixed effects and facility characteristics including chain ownership status, demographic characteristics of the zip code, and whether the facility is freestanding or hospital based. The dependent variable in col. 1 is an indicator for whether the patient rides in the following month. The dependent variable in col. 2 is an indicator for whether the patient is hospitalized in the same month in which they are observed to be riding. The dependent variable in col. 3 is an indicator for whether a patient dies in the same month that they are observed to be riding. Sample is limited to patient-months in which the patient receives at least one nonemergency dialysis ambulance ride. Standard errors clustered at the district level are given in parentheses.

+ Significance at the 10% level.

* Significance at the 5% level.

ambulance rides after the start of prior authorization are less healthy, which is consistent with Medicare's aim for the program: to provide rides only when medically necessary.

The denial rate for submitted claims provides additional evidence that prior authorization resulted in a more appropriate use of ambulances.

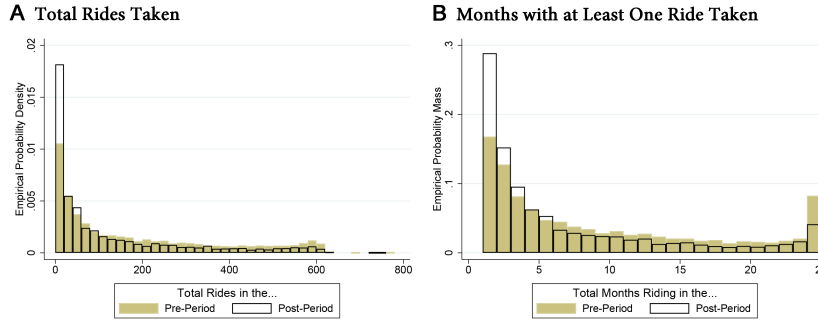

FIG. 8.—Histogram of ridership among riders. *A*, Histograms of total rides taken by patients in districts subject to prior authorization in the 24 months before and after the implementation of prior authorization. *B*, Analogous histograms for the total number of months in which the patient takes at least one ride. All rides are nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data.

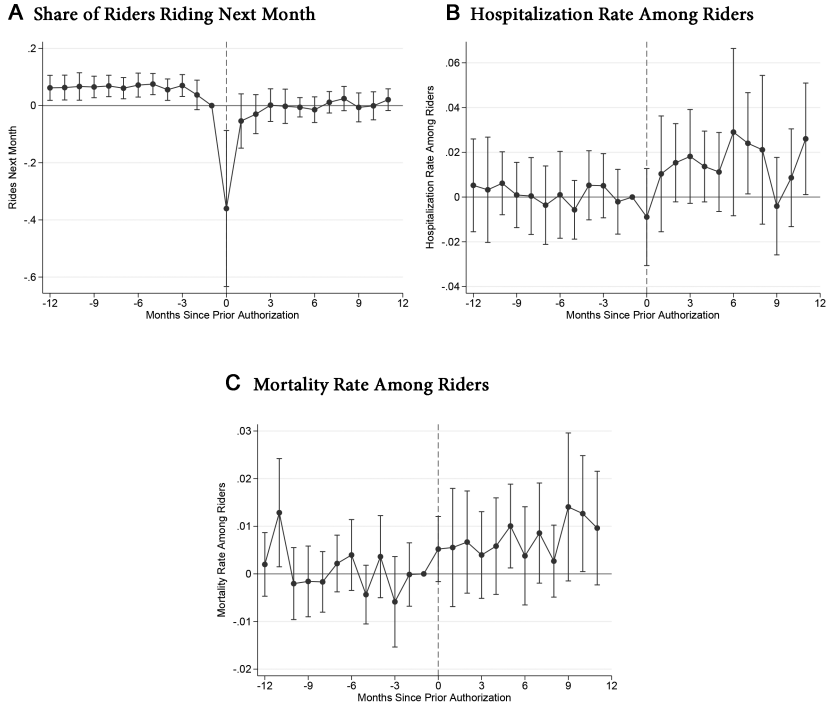

FIG. 9.—Effect of prior authorization on patient selection: estimates of β_e for $e \in [-12, 11]/\{-1\}$ from equation (3). These data include rides from 2011 to 2017. An observation is a patient-month. Controls include incident patient characteristics, age, and tenure on dialysis as well as facility fixed effects and facility characteristics including chain ownership status, demographic characteristics of the zip code, and whether the facility is freestanding or hospital based. Sample is limited to patient-months in which the patient receives at least one nonemergency dialysis ambulance ride. Standard errors are clustered at the district level. Error bars represent pointwise 95% confidence intervals.

Although we do not observe the requests submitted by providers to obtain prior authorization, we do observe whether a claim was paid after it was submitted for reimbursement. Figure 10 shows that, immediately following prior authorization, the share of claims denied by Medicare jumped sharply and then declined gradually.[18] Furthermore, figure A12 shows the denial rates of all firms in panel A contrasted with the denial rates of only those firms that continued providing nonemergency rides in

[18] Because these denial rates capture only claims that were submitted after providers could obtain prior authorization for the service, rather than including those that were denied prior authorization, the increase in denial rates after prior authorization is likely a lower bound for the true increase. Indeed, the CMS (2020a) reports that in the first year of prior authorization only 35% of prior authorization requests were affirmed while in subsequent years this number was between 57% and 66%.

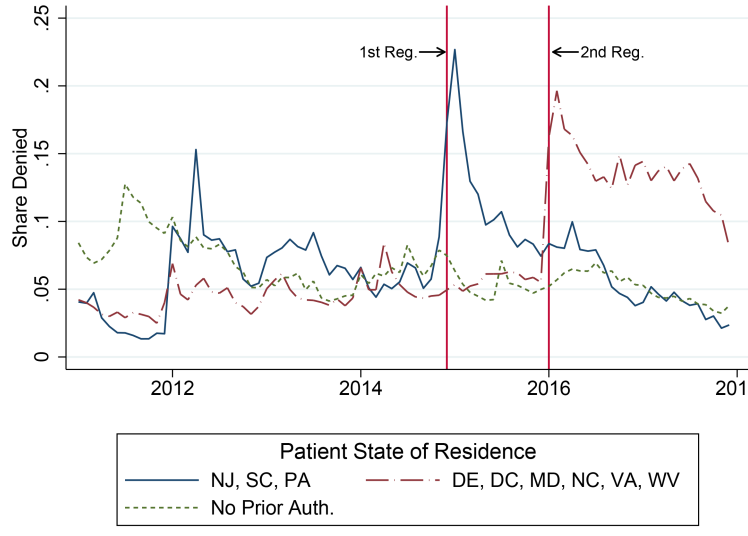

FIG. 10.—Claim denial rates by prior authorization status. The sample includes nonemergency basic life-support ambulance rides from a dialysis facility to a place of residence for ESRD patients from 2011 to 2017. State is determined by the transported patient's state of residence. Vertical lines mark the implementation of prior authorization in New Jersey, South Carolina, and Pennsylvania, and in Delaware, Washington, D.C., Maryland, North Carolina, Virginia, and West Virginia. The share of claims denied is the share of rides for which the submitted claim was not paid any positive amount.

panel B. Both panels have similar patterns for denial rates, although the spike is slightly less pronounced for those that continued to serve the market. In panel C, we decompose the sample further into firms exiting in the first months of prior authorization, those that did not exit immediately but that did not continue providing rides for at least the next 2 years, and those that continued regularly providing rides; the spike in denials is most pronounced for firms that exited immediately upon the start of prior authorization. These results suggest that the pattern in denial rates comes from both firms whose claims were denied and then exited the market as well as those whose denial rates initially increased and then declined. That the overall denial rate decreased following the initial spike indicates that some firms stopped submitting claims that would be denied under the heightened scrutiny of prior authorization, which we interpret as evidence that prior authorization acts as a screening mechanism that effectively deters fraud.

D. Mechanisms of Litigation

In line with its limited impact on ridership and payments, we also find that realized litigation had a negligible effect on the overall structure

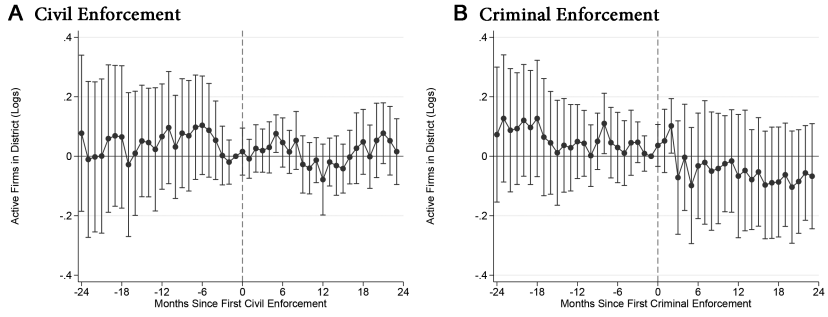

FIG. 11.—Effect of litigation on number of firms in district: estimates of β_e for $e \in [-24, 23]/\{-1\}$ from equation (1). Dependent variable is the number of firms providing nonemergency basic life-support rides between a dialysis facility and a patient's home in a district-month transformed by adding 1 and taking the natural log. These data come from a 20% sample of all Medicare beneficiaries and include rides from 2007 to 2019. An observation is a district-month. Standard errors are clustered at the district level. Error bars represent pointwise 95% confidence intervals.

of the market for ambulance companies. Both figure 11 and table 8 demonstrate that civil enforcement does not reduce the number of active firms, while criminal enforcement leads to an imprecisely estimated 4.3% drop. In appendix G, we further show that, unlike prior authorization, litigation does not affect firm specialization.

In contrast to its impact on the wider market, we do find that criminal litigation effectively incapacitates the prosecuted firms themselves. Figure 12 uses our 20% sample of all Medicare beneficiaries' rides to show that payments fall nearly to zero for firms subject to criminal indictments shortly following the indictment, whereas civil litigation has no apparent effect on the indicted firms.

TABLE 8
EFFECT OF LITIGATION ON NUMBER OF ACTIVE FIRMS

	Civil		Criminal	
	Active Firms (Log) (1)	Active Firms (2)	Active Firms (Log) (3)	Active Firms (4)
Enforcement	.0122	.779	−.0442	−5.651
	(.0290)	(.652)	(.0731)	(6.436)
Month-year fixed effects	1	1	1	1
District fixed effects	1	1	1	1
Dependent variable mean	1.298	7.192	1.314	6.906
Observations	12,143	12,143	12,203	12,203

NOTE.—Estimates of β from eq. (2). Dependent variables are the number of firms providing nonemergency basic life-support rides between a dialysis facility and a patient's home in a district-month and the natural logarithm of 1 plus the same. These data come from a 20% sample of all Medicare beneficiaries and include rides from 2007 to 2019. An observation is a district-month. Standard errors are clustered at the district level.

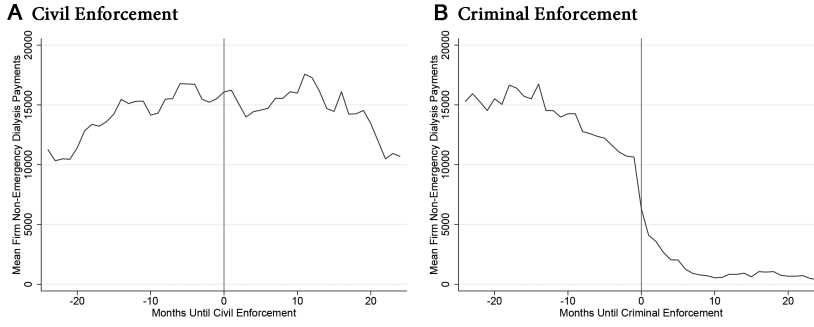

FIG. 12.—Estimates of incapacitation effect. Average monthly Medicare payments to firms subject to civil or criminal enforcement in the 24 months before and after complaint or indictment date. These data come from a 20% sample of all Medicare beneficiaries and include rides from 2007 to 2019. An observation is a firm-month.

We can further disentangle the respective mechanisms of litigation by contrasting an incapacitation effect—the direct effect of an enforcement action on the defendants themselves—against a deterrence effect of litigation on the other firms in the market not included in the lawsuit. The approximately $12,000 per month reduction in payments per firm following criminal indictments scales up to approximately $60,000 when accounting for our 20% sample of claims. Moreover, in districts subject to criminal litigation in the period for which we observe firm identifiers, the DOJ indicted 14 firms across all districts in the 2 years following the first indictment, which means that the estimated treatment effect of criminal litigation overall corresponds to the effect of 1.6 firms being indicted, on average, in each district. Based on these estimates, only $93,000 of the $615,000 per district-month reduction in spending in table A13 comes from indicted firms, with the remainder coming from firms not directly tied to the enforcement action. Put differently, the criminal incapacitation effect accounts for a comparatively small 15.2% of the overall effect of realized criminal enforcement. Compared with prior authorization, where the regulation reduces payments and rides through claim denials and, consequently, drives many firms out of the market, litigation has both a direct incapacitation effect on the firms indicted in the lawsuit and a larger deterrence effect on the firms not being indicted as they learn about the enforcement action and decide whether to change their behavior in response.

To test whether lawsuits have a limited impact because firms respond to the threat of enforcement, called general deterrence, rather than realized enforcement, we use hand-collected data from the Department of Justice on the personnel hours devoted to civil and criminal enforcement in each federal court district and measure enforcement capacity at a district-year level by the number of hours spent in federal criminal or civil court

TABLE 9
EFFECT OF ENFORCEMENT CAPACITY ON AMBULANCE SPENDING AND RIDES

	Total Ride Payments (Log) (1)	Total Rides (Log) (2)	Total Ride Payments (Log) (3)	Total Rides (Log) (4)	Total Ride Payments (Log) (5)	Total Rides (Log) (6)
Civil court hours (log)	.0117	−.0124				
	(.105)	(.0574)				
Criminal court hours (log)			.0631	−.0705		
			(.191)	(.144)		
Total court hours (log)					.102	−.0478
					(.262)	(.196)
Month-year fixed effects	1	1	1	1	1	1
District fixed effects	1	1	1	1	1	1
Dependent variable mean	12.74	7.701	12.74	7.701	12.74	7.701
Observations	1,410	1,410	1,410	1,410	1,410	1,410

NOTE.—Estimates from a regression of measures of ridership on log personnel hours: civil hours in cols. 1 and 2, criminal hours in cols. 3 and 4, and total hours in cols. 5 and 6. All rides are nonemergency basic life-support rides between a dialysis facility and a patient's home observed in the USRDS data. Dependent variables are transformed by adding 1 and taking the natural log. These data include rides from 2003 to 2017. An observation is a district-month. All specifications include district and time fixed effects. Standard errors are clustered at the district level.

by attorneys in the US Attorney's Office in that district-year on all types of cases, not just health-care fraud.[19]

Table 9 shows the effect of various types of enforcement capacity on rides and payments. We do not find a meaningful relationship between ridership and personnel hours for any measure, ruling out at the 95% confidence level, for example, an elasticity of payments with respect to civil enforcement capacity of −0.20 and an elasticity with respect to criminal capacity of −0.32. Although actively pursuing criminal litigation can reduce spending, we see no clear effect of marginal changes in latent enforcement capacity by itself, indicating that there is no general deterrence effect from an increased risk of prosecution.

VI. Why Up-Front Regulation Outperformed Pay-and-Chase

An extensive theoretical literature has considered whether ex ante regulation or ex post litigation is more effective at combating illegal

[19] See https://www.justice.gov/usao/resources/annual-statistical-reports.

behavior. Much of this prior work has addressed torts and property rights violations, where individuals or private parties are harmed. We provide an important and natural extension of these studies to circumstances where the injured party is the government, the type of crime is financial fraud, and the illegal behavior is perpetrated by a large number of fraudulent actors. To frame these empirical results, we develop a stylized model to revisit the question of when and how litigation may effectively deter fraud on its own or when regulation must be used in conjunction with it.

Consider a firm deciding whether to commit fraud. The firm will do so if

$$G(Reg) > P_{\text{Crim}}F_{\text{Crim}} + P_{\text{Civ}}F_{\text{Civ}}, \tag{5}$$

where $G(Reg)$ is the gain from fraud, which depends on whether prior authorization is in place, captured by $Reg \in \{0, 1\}$; P_{Crim} and P_{Civ} are the perceived probabilities of facing criminal or civil enforcement; and F_{Crim} and F_{Civ} are the criminal and civil penalties the firm faces if caught and successfully prosecuted. The gains from fraud are the difference between the fraudulent payments the firm receives from Medicare and the firm's operating costs,

$$G(Reg) = R(Reg) - C(Reg).$$

Costs $C(Reg)$ are higher under prior authorization due to the hassle costs of navigating the regulatory environment, while revenue $R(Reg)$ is lower because the regulation leads to more claim denials, reducing the ability of the firm to steal funds in the first place. The penalties for being caught are

$$F_{\text{Civ}} = \min(3R(Reg), Assets) \quad \text{and} \quad F_{\text{Crim}} = \min(3R(Reg), Assets) + J,$$

which reflects the stipulation of the False Claims Act that financial penalties from a civil judgment are three times the amount stolen but bounded by the firm's assets. That is, the firm faces only limited liability. The parameter J within criminal enforcement captures the firm operator's disutility from going to jail, and jail costs can be imposed in criminal cases even against firms unable to pay the financial penalty.

This stylized model highlights the main factors that determine the relative effectiveness of civil litigation, criminal litigation, and prior authorization and provides a framework for explaining both why firms committed fraud and why regulation was much more effective at stopping them. In short, the potential for lucrative Medicare reimbursements coupled with firms' limited liability and low probability of being detected resulted in widespread fraudulent activity. Prior authorization helps resolve these issues by limiting the initial gains from committing fraud while at the

same time not requiring high probabilities of detection or high rates of recovery.

A. *Limited Liability*

In the case of ambulance fraud, the government faces several constraints that make litigation unlikely to have a widespread effect on illegal behavior. First among these is firms' limited liability, as litigation may fail to curtail illicit behavior if severe penalties cannot be enforced (Shavell 1984; Polinsky and Shavell 2000). A fly-by-night ambulance company can spend its ill-gotten gains G before being prosecuted and can shut down in response to the financial penalties imposed by the courts, making a pay-and-chase approach largely ineffective. In the model, this is captured in $min(3R, Assets)$, where *Assets* are endogenously chosen by the firm and may be drawn down quickly to reduce the amount that might have to be repaid upon conviction. Even for a successfully prosecuted firm, the state's likelihood of securing full restitution is low, essentially limiting the firm's liability. Despite judgments regularly reaching millions of dollars, the DOJ warns that restitution for criminal penalties is often difficult to enforce, writing, "Realistically, however, the chance of full recovery is very low . . . it is rare that defendants are able to fully pay the entire restitution amount owed" (DOJ 2021).

To test this hypothesis in our empirical setting, we filed Freedom of Information Act requests with each of the US Attorney's Offices for the actual financial recoveries from all of the ambulance fraud cases involving dialysis patients in which we observe a prosecution. We were able to determine recovery amounts for 27 cases, which averaged less than \$1.2 million in recovered funds, or 51% of the total amount owed. In only 10 of these cases has the full amount of penalties been paid, while in 11 cases the recovery funds amount to less than 20%, reflecting the limited liability of many defendants. Given that the median case for which we have data closed in 2016, it seems unlikely that the full amount for the remaining cases will ultimately be recovered.

The challenge of enforcing financial penalties against this particular population may explain why criminal lawsuits are more effective than civil enforcement. Civil lawsuits impose only monetary penalties or exclusion from the Medicare program, penalties that may not have much impact on firms that can either shut down after allegations of fraud or even continue committing fraud after paying a fine. Conversely, criminal lawsuits can impose jail time on the owners or operators of fraudulent firms, a nonmonetary penalty that can be enforced even in the absence of recoverable funds and incapacitate the operator. This is reflected in our model by J, which is not subject to limited liability. In section V and appendix I, we show that, in practice, nearly all of the accused firms in our data shut

down after criminal indictments, whereas civil complaints had almost no effect on the probability of the targeted firm remaining in the market.

Limited liability is not confined to firms alone, as the beneficiaries who participate in the fraud may escape liability as well. Although not included in our stylized model, patients were often a key part of the fraudulent schemes, with some criminal lawsuits alleging that they received kickbacks for riding and referring others. We identify 6,789 unique beneficiaries who rode with the firms that were prosecuted from 2012 to 2017 and more than 2,700 who immediately stopped riding in the first three states subject to prior authorization, perhaps reflecting a large faction of complicit beneficiaries. Despite compelling evidence of widespread involvement among dialysis patients, the government has criminally prosecuted only four of them for health-care fraud, likely owing to their vulnerable conditions as well as the exorbitant costs of imprisoning them in one of the six overcrowded Bureau of Prisons Medical Centers, the institutions for prisoners with acute medical needs like dialysis (DOJ 2015; BOP 2019).

B. Low Probability of Detection

In addition to the challenge of levying and collecting large penalties against fraudulent firms, litigation may also be hindered by the difficulty of detecting and successfully prosecuting illicit behavior at a sufficiently large scale. In our model, the perceived probability of enforcement is captured by P_{civ} and P_{crim}, where lower values would mean that firms have a higher expected value of committing fraud. From 2007 to 2014, only 28 firms were subject to criminal litigation and 44 to civil litigation, while we estimate that approximately 1,150 firms may have provided fraudulent rides over this period, implying just a 2.4% chance of being pursued for criminal litigation and 3.8% for civil.[20] The perceived probability of enforcement for sophisticated firms should center around these levels, although firms may update their beliefs after seeing actual enforcement in their districts, which could explain the modest effect criminal enforcement has on firms in our setting.

Relatedly, firms do not appear to respond to the threat of litigation due to overall enforcement capacity. As we show in table 9, increases in the underlying enforcement capacity of the DOJ do not change firm behavior. Interpreted through the lens of our model, this could be because firms do not update their perceived probability of detection due to changes in overall enforcement capacity.

One primary reason for such low detection rates is that health-care fraud can be difficult to prove after the fact and criminal lawsuits require

[20] Details on these calculations are available in app. J.

a "beyond a reasonable doubt" evidentiary standard, such as video recordings of purportedly bedridden patients walking on their own.[21] With thousands of firms providing nonemergency ambulance rides and the limited resources of the DOJ and FBI constraining their ability to widely prosecute such cases, the chance that any given fraudulent ambulance company will be detected is very low.

A lack of specialization among prosecutors and judges may also partly explain the low detection rates (Landis 1938). Almost two dozen different judicial districts were involved in the lawsuits that we study, which means that dozens of different investigators, attorneys, and judges were responsible for understanding the complex nature of this fraud in order to successfully prosecute it. Moreover, DOJ attorneys who work on health-care fraud are responsible for enforcing many other parts of the federal criminal and civil code, as are the judges who try the cases. We validate this empirically with data from the DOJ National Caseload Data from 2001 to 2021. Among the US Attorney's Office staff ever assigned as lead attorney in a health-care fraud case, the median attorney has five criminal health-care fraud cases throughout their career, constituting only 1.8% of their case load, with an interquartile range of 0.4%–7.9%.

The low rates of detection for ambulance fraud relate to the work of Behrer et al. (2021) and Mookherjee and Png (1992), who argue that, in the case of private harms, litigation alone is ineffective when the harm in question affects a large number of individuals and the private reporting of harm is insufficient. In the case of health-care fraud, the injured party is every US taxpayer, and individuals are not empowered to protect the public interest. The government also faces agency costs, because the stolen money does not directly impact the federal employee charged with carrying out enforcement. That is, failing to detect health-care fraud has limited consequences for those directly responsible for combating it.

C. Gains from Fraud

While the risks to fraudulent firms from litigation were low, the potential gains were large before the onset of prior authorization, as reflected by $G(Reg = 0)$ in our model. Among the 65 litigated firms that we observe in our claims data from 2007 to 2019, each received an estimated $5.4 million in payments from Medicare.[22] Although we do not have precise estimates of the costs of perpetrating this fraud, anecdotal evidence suggests they are very low.

[21] For example, such evidence was used in the prosecution of Saltville Rescue Squad; case 1:12-00002, Western District of Virginia.

[22] This number is approximate because we observe only a 20% sample of claims for these firms.

Given the low probability of detection—and, conditional on being prosecuted, the limited recovery rate of fraudulent payments—the expected financial cost of fraud is approximately $72,000.[23] Ignoring jail time, this figure implies that committing fraud is profitable as long as the firm has a profit margin greater than 1.4%, an exceedingly low hurdle that can explain the widespread proliferation of ambulance taxis before prior authorization.

Appendix J presents details of a broader calibration exercise, where we use actual recoveries to estimate penalties subject to limited liability, lawsuit counts to estimate probabilities of enforcement, and claims to estimate revenue gains from fraud. The calibrated estimates show that, due to limited liability, even a civil enforcement probability of one would not be as large a deterrent as prior authorization given the limited liability firms face from civil enforcement.

D. Why Regulation Succeeded

Regulation succeeded where litigation failed because it solves the problems of limited liability and a low probability of detection by directly reducing the potential gains from fraud. In the context of our model, this would be captured by a low value for $G(Reg = 1)$, so even with low probabilities of detection and relatively small fines, fraud is no longer profitable. In short, regulation succeeds primarily by preventing fraudulent funds from being paid out in the first place. Under prior authorization, firms are not paid for their claims until they establish their patients meet Medicare's criteria for a medically necessary ride, which means they cannot spend their ill-gotten gains during the intervening period when the fraud goes undetected. In the model, this comes from a sharp reduction in $R(Reg)$. As shown by figure 10, the claim denial rate spiked in states subject to prior authorization, going from 5.7% in the year before prior authorization to 22.7% in January 2015. Limiting the sample to firms that exited at the start of prior authorization, and who are therefore more likely to be fraudulent, we find that the denial rate jumped from 8.1% to 52.5%. Furthermore, the Centers for Medicare and Medicaid Services reports that 65% of prior authorization requests submitted in 2015 were denied, with this rejection rate falling to 44% over the next three years (CMS 2020a). Based on these figures, prior authorization reduced fraudulent firm revenues by roughly 70%, or $3.8 million per firm on average.

Beyond significantly curtailing the ability of fraudulent firms to extract revenue from Medicare, prior authorization also increases firms' costs

[23] This figure is the estimated probability of facing civil or criminal litigation multiplied by the average amount recovered by the government in the cases we observe. See app. J for details.

through administrative burdens and paperwork, captured by $C(Reg = 1)$. Although others have estimated these hassle costs to be large in some settings (e.g., Herd and Moynihan 2018; Dunn et al. 2023; League 2023), we consider it unlikely that prior authorization imposes a large burden on patients or physicians in this case. For example, CMS focus groups of physicians indicated that nursing staff generally fill out the forms before a physician signs them, likely imposing a low cost of complying with the regulation. Although ambulance companies expressed more frustration with the process, denials of prior authorization "typically resulted from beneficiaries not meeting CMS's existing (premodel) medical necessity requirements" rather than from clerical errors in filling out the proper paperwork (Weinstock et al. 2018). Furthermore, calibrating the paperwork cost of prior authorization for ambulance rides to those found elsewhere in the literature, we estimate that, even under extreme assumptions, they amount to no more than $3,500 over the entire life of the average ambulance company.[24] Increasing the costs of fraud appears to have contributed much less to the effectiveness of prior authorization than reducing the initial outlay of revenue did.

The large drop in revenue paired with the very modest increase in costs stemming from prior authorization rationalizes the large reduction in fraud that we observe in the data. Using our estimate on the revenue-reducing effect of prior authorization, our model indicates that even if we ignore the disutility of jail time, fraud would be unprofitable under prior authorization as long as fraudulent firms' profit margins without prior authorization are less than 250%.[25] The primary advantage that prior authorization has over litigation, then, is its ability to prevent improper payments from ever being paid out.

Beyond the framework of our stylized model, regulation may complement litigation in other important ways as well. For example, regulations may improve detection rates by making noncompliance more obvious and easier to prosecute in court. Although courts may find it difficult to assess medical necessity, regulations can create "bright-line rules" that

[24] Previous estimates of the cost of a prior authorization request range up to just over $30 per request, or almost $90 per successful request using the published rate of request affirmation (CMS 2020a). Scaling this amount by the average number of patients, we observe a firm serve over its life (approximately 39), we arrive at a total cost of just under $3,500. See app. J for details.

[25] We estimate the average fraudulent firm's total revenue is approximately $5.4 million, the total additional paperwork cost of prior authorization is roughly $3,500, and the expected monetary cost of detection is approximately $72,000. Because we estimate that roughly 70% of claims are denied under prior authorization (due to either the prior authorization request or subsequent claim being denied), the nonprior authorization costs of operating a fraudulent firm must be less than $1.5 million to make operating under prior authorization profitable. This implies a minimum necessary profit margin of ($5.4 million − $1.5 million)/$1.5 million, or roughly 250%. See app. J for details.

are easy to monitor (Kaplow 1992; Glaeser and Shleifer 2002). With prior authorization, it is much simpler to provide enough evidence that a firm failed to submit paperwork than it is to prove that a patient did not have a legitimate medical reason for using an ambulance. As discussed in Glaeser and Shleifer (2001), simple, easy-to-enforce regulations strengthen the ability of the government to stop illegal behavior.

Also related is the prior theoretical work of Glaeser and Shleifer (2003) comparing pure litigation-based enforcement to a regime that uses administrative rules as well. Most relevant for our setting, they find that adding administrative rules is optimal in cases where litigation can be subverted. Although not addressed in prior work, the unwillingness of prosecutors to pursue complicit beneficiaries and the challenge of recovering funds from fly-by-night firms are both forms of subversion that make litigation ineffective at assigning liability on its own. We provide suggestive evidence in appendix I that prior authorization was especially effective at shutting down what appear to be fly-by-night firms, as the increased likelihood that a firm exits after prior authorization was most pronounced among small firms that specialized almost entirely in nonemergency ambulance services.

Administrative enforcers can also be more specialized than judges or prosecutors, which facilitates enforcement (Landis 1938). As we discussed above, DOJ attorneys are not medical experts or even specialists in healthcare fraud; these attorneys must convince unspecialized judges and juries that care was not medically necessary, a challenge perhaps best reflected by the ongoing circuit split, in which different appellate courts have different standards for whether medical decisions without a consensus opinion can be prosecuted for fraud (Jones Day 2021). By contrast, the administrators responsible for checking prior authorization requirements for ambulance reimbursements focus solely on Medicare regulations and are well equipped to evaluate medical necessity.[26]

E. Relative Cost-Effectiveness

By modeling the decision of firms to commit fraud, our model allows us to understand the effectiveness of regulation and litigation rather than the desirability of these policies. Although a full model of welfare is beyond the scope of the current paper, in this section we discuss the

[26] As one administrator noted, "The staff reviewing these claims will be experienced with Medicare's coverage, coding and payment requirements for existing policies and procedures" (Mauch, personal communication, 2022), while another emphasized that "clinical reviewers receive specialized training for the types of services they are reviewing and have detailed procedures to reference for consistent, calibrated review approaches" (Portzline, personal communication, 2022).

key consideration beyond effectiveness for which we have the best data: the financial costs of implementing the policies.[27]

Because monitoring paperwork for prior authorization is much simpler than conducting ex post enforcement against fraudulent claims, regulation can accomplish a higher level of deterrence at a much lower cost. As it relates to our setting, the chief actuary for CMS estimated the cost of implementing prior authorization nationwide at only "$38.1 million in the first expansion year and $28.6 million per year in subsequent years" (Spitalnic 2018). Given that we estimate a reduction in Medicare spending of more than $300 million in the eight states subject to the pilot program in its first 2 years, prior authorization is much more cost-effective than widespread litigation at $250,000–$300,000 per case.[28] Our results in table A20 suggest that the $6,500,000 spent on litigating civil cases (26 cases × $250,000/case) had no effect beyond the cases themselves, and table A20 suggests that prosecuting additional firms does not amplify the effects of criminal litigation. In the context of the model, realized criminal and civil litigation can potentially reduce fraud among nonindicted firms by raising the perceived probability of detection, either because the actual detection probabilities are higher or because their salience increases subjective beliefs about them. Even when a fraudulent ambulance company faces civil prosecution with absolute certainty, however, our model estimates still imply that prior authorization has greater deterrence effects in light of firms' limited liability. Expending additional resources pursuing widespread litigation will never be able to achieve the deterrence of prior authorization and would cost a great deal more.

In addition to deterrence, regulation and litigation can have other effects that are difficult to measure empirically. In response to the increased scrutiny of ambulance taxis, some firms may simply choose to forgo this

[27] A full welfare model would require specifying the social planner's objective function and measuring parameters for which we do not have reasonable estimates, e.g., social welfare weights associated with different parties' utilities or profits, the outside option of the courts and social (dis)utility of jail time that arise from litigation, and the weight on the firms' limited paperwork costs as well as the disutility from firms' giving back or forgoing illicit profits.

[28] We arrive at this estimate using two different approaches. First, Leder-Luis (forthcoming) measures public spending on False Claims Act cases, finding $108.5 million spent on 446 civil cases, or $243,000 per case. Second, the Federal Health Care Fraud and Abuse Control Program Annual Report provides details on the number of civil and criminal health-care fraud investigations, estimating that $1,059,315,473 was spent on 3,603 investigations in 2019, or $294,000 per case. Specifically, the DOJ opened 1,060 new criminal health-care fraud investigations, and it opened 1,112 new civil health-care fraud investigations. In addition, investigations conducted by HHS-OIG resulted in 747 criminal actions and 684 civil actions against individuals or entities that engaged in crimes related to Medicare and Medicaid. We arrive at the 3,603 figure by summing these investigations and actions. Note that our estimate is likely somewhat biased downward as an estimate of the cost of litigation since there are investigations that do not result in actions, and these figures do not include other relevant budgetary figures (e.g., from the FBI budget).

type of fraudulent activity in the first place, a general deterrence effect of unknown magnitude (Shavell 1991; Leder-Luis, forthcoming). Conversely, individuals intent on committing health-care fraud may substitute away from one particular scheme and pursue others that are more difficult for authorities to detect. On the other hand, regulation may create additional nonmonetary costs, such as if it results in care being rationed inefficiently (AMA 2021), which then leads to a lower quality of care. As noted above, however, we find no evidence that prior authorization led to worse outcomes for patients in our setting.

Finally, the administrative burden associated with regulation may impose hassle costs on nonfraudulent firms or may benefit payers by serving as a screening mechanism. As discussed above, the paperwork costs of this particular regulation are low, particularly when compared with the reduction in Medicare spending. As shown in table A24, even under extreme assumptions well outside the range of cost estimates in the literature, the total paperwork costs of prior authorization are less than $60,000 per district per month, or only 8.1% of the estimated reduction in Medicare spending. Beyond its relatively low direct costs, regulation may be well targeted such that only medically necessary services are rendered, as providers and patients anticipate that only valid claims will be approved (Zeckhauser 2021), resulting in an equilibrium in which the regulation is not costly to enforce because fewer claims are filed in the first place. In our setting, this is consistent with the changes that we observe for both denial rates and the mix of patients riding an ambulance following prior authorization.

VII. Conclusion

We find that imposing prior authorization on ambulance rides for dialysis patients was much more effective at reducing wasteful spending than pursuing criminal or civil litigation on their own. Prior authorization caused an immediate and persistent drop in nonemergency ambulance rides of nearly 68%, whereas lawsuits against fraudulent providers had a much smaller effect. Had the federal government required prior authorization throughout our sample period, it would have saved $4.8 billion and prevented 21.2 million unnecessary rides at an administrative cost of only $28 million per year (Spitalnic 2018).[29] When compared with pay-and-chase enforcement and the relatively large costs associated with it, prior authorization is much more efficient.

Importantly, we show that the decrease in nonemergency rides did not come at the expense of patients' health, even though it drove many

[29] See app. K for details of the calculation of savings from prior authorization.

ambulance companies out of the market. Following prior authorization, patients who continued taking nonemergency ambulance rides to their dialysis sessions were in poorer health, suggesting that the benefit was being used more efficiently and as intended by Medicare.

Our results relate to the economic theory of why regulation is necessary—and litigation alone insufficient—for successfully combating Medicare fraud. Criminal and civil penalties are often too low given prosecutors' inability to levy large penalties against fly-by-night firms, and prosecution rates are held back by the challenges of detecting fraud, the diffuse nature of the harm, and the limited resources of unspecialized enforcers. This points to health-care fraud as being an area in need of regulatory innovations to complement the use of legal enforcement through prosecution. Medicare has recently moved in this direction, expanding prior authorization to other medical expenditures that may be especially susceptible to fraud, such as power mobility devices, home health services, and hyperbaric oxygen. Our results suggest that such reforms are likely to be successful.

Our results also highlight a way to reduce fraud in other areas of government expenditure. Whenever dealing with a multitude of small firms, the government faces the same challenges of limited liability and a low probability of detection that hindered its response to ambulance taxis. In cases such as pandemic aid (Griffin, Kruger, and Mahajan 2021; Autor et al. 2022) and defense contracting (Karpoff, Lee, and Vendrzyk 1999), regulations like prior authorization that verify up front whether a payment is appropriate can be used to deter fraud effectively.

Data Availability

Eliason et al. (2024) in the Harvard Dataverse, https://doi.org/10.7910/DVN/QAGBDM, contains publicly available data on litigation activity, DOJ and US attorney workloads, and publicly available Medicare data. It also includes instructions on accessing the proprietary claims data used in this analysis. Finally, it contains code for replicating the analysis along with result and log files.

References

Ackerman, A., and A. Omeokwe. 2022. "Covid-19 Relief Fraud Potentially Totals $100 Billion, Secret Service Says." *Wall Street J.*, Dec. 22.

Agan, A. Y., J. Doleac, and A. Harvey. 2023. "Misdemeanor Prosecution." *Q.J.E.* 138 (3): 1453–505.

Agan, A. Y., M. Freedman, and E. Owens. 2021. "Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense." *Rev. Econ. and Statis.* 103 (2): 294–309.

AMA (American Medical Association). 2021. "2020 AMA Prior Authorization (PA) Physician Survey." Technical report, AMA, Chicago.

Autor, D., D. Cho, L. D. Crane, M. Goldar, B. Lutz, J. K. Montes, W. B. Peterman, D. D. Ratner, D. Villar Vallenas, and A. Yildirmaz. 2022. "The $800 Billion Paycheck Protection Program: Where Did the Money Go and Why Did it Go There?" Working Paper no. 29669, NBER, Cambridge, MA.

Becker, G. S. 1968. "Crime and Punishment: An Economic Approach." *J.P.E.* 76 (2): 169–217.

Behrer, A. P., E. L. Glaeser, G. A. M. Ponzetto, and A. Shleifer. 2021. "Securing Property Rights." *J.P.E.* 129 (4): 1157–92.

Boning, W. C., J. Guyton, R. Hodge, and J. Slemrod. 2020. "Heard It through the Grapevine: The Direct and Network Effects of a Tax Enforcement Field Experiment on Firms." *J. Public Econ.* 190:104261.

BOP (Federal Bureau of Prisons). 2019. "Care Level Classification for Medical and Mental Health Conditions or Disabilities." https://www.bop.gov/resources/pdfs/care_level_classification_guide.pdf.

Borusyak, K., X. Jaravel, and J. Spiess. 2017. "Revisiting Event Study Designs." Scholarly Paper no. 2826228, SSRN, Rochester, NY.

Bowen, S. 2013. "Learning from Iraq: A Final Report from the Special Inspector General for Iraq Reconstruction." Council on Foreign Relations, Bethesda, MD.

Brot-Goldberg, Z., S. Burn, T. Layton, and B. Vabson. 2022. "Rationing Medicine through Bureaucracy: Authorization Restrictions in Medicare." Working Paper no. 30878, NBER, Cambridge, MA.

Callaway, B., and P. H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *J. Econometrics* 225 (2): 200–230.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs." *Q.J.E.* 134 (3): 1405–54.

CMS (Centers for Medicare and Medicaid Services). 2020a. "Medicare Prior Authorization Model for Repetitive, Scheduled Non-Emergent Ambulance Transports Status Update." https://www.cms.gov/files/document/status-update-11-18-2020.pdf.

———. 2020b. "Repetitive, Scheduled Non-Emergent Ambulance Transport (Rsnat) Prior Authorization Model Frequently Asked Questions." https://www.cms.gov/research-statistics-data-and-systems/monitoring-programs/medicare-ffs-compliance-programs/prior-authorization-initiatives/downloads/ambulancepriorauthorization_externalfaq_121517.pdf.

———. 2023. "Fiscal Year 2023 Improper Payments Fact Sheet." https://www.cms.gov/newsroom/fact-sheets/fiscal-year-2023-improper-payments-fact-sheet.

Coase, R. H. 1960. "The Problem of Social Cost." *J. Law and Econ.* 3 (2): 1–44.

Contreary, K., A. Asher, and J. Coopersmith. 2022. "Evaluation of Prior Authorization in Medicare Nonemergent Ambulance Transport." *JAMA Health Forum* 3 (7): e222093.

Dafny, L. S. 2005. "How Do Hospitals Respond to Price Changes?" *A.E.R.* 95 (5): 1525–47.

DOJ (Department of Justice). 2015. *The Impact of an Aging Inmate Population on the Federal Bureau of Prisons*. Justice Department, Office of the Inspector General.

———. 2021. "Restitution Process." https://www.justice.gov/criminal/criminal-vns/restitution-process#:~:text=Restitution%20payments%20received%20by%20the,receives%20payments%20from%20the%20defendant.

Dunn, A., J. D. Gottlieb, A. H. Shapiro, D. J. Sonnenstuhl, and P. Tebaldi. 2023. "A Denial a Day Keeps the Doctor Away." *Q.J.E.* 139 (1): 187–283.

Eliason, P., R. League, J. Leder-Luis, R. McDevitt, and J. Roberts. 2024. Replication Data for: "Ambulance Taxis: The Impact of Regulation and Litigation on

Health Care Fraud." https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/QAGBDM.
Eliason, P. J., B. Heebsh, R. C. McDevitt, and J. W. Roberts. 2020. "How Acquisitions Affect Firm Behavior and Performance: Evidence from the Dialysis Industry." *Q.J.E.* 1 (135): 221–67.
Esson, M. I. 2021. "Incentive Misalignment in Ambulance Reimbursement: An Analysis of Medicare Payment Systems." Working paper.
Fang, H., and Q. Gong. 2017. "Detecting Potential Overbilling in Medicare Reimbursement via Hours Worked." *A.E.R.* 107 (2): 562–91.
———. 2020. "Detecting Potential Overbilling in Medicare Reimbursement via Hours Worked: Reply." *A.E.R.* 110 (12): 4004–10.
Glaeser, E. L., and A. Shleifer. 2001. "A Reason for Quantity Regulation." *A.E.R.* 91 (2): 431–35.
———. 2002. "Legal Origins." *Q.J.E.* 117 (4): 1193–229.
———. 2003. "The Rise of the Regulatory State." *J. Econ. Literature* 41 (2): 401–25.
Griffin, J. M., S. Kruger, and P. Mahajan. 2021. "Did FinTech Lenders Facilitate Paycheck Protection Program Fraud?" Scholarly Paper no. 3906395, SSRN, Rochester, NY.
Gruber, J., D. H. Howard, J. Leder-Luis, and T. L. Caputi. 2025. "Dying or Lying? For-Profit Hospices and End of Life Care." *A.E.R.* 115 (1): 2637–94.
Harrington, C., J. Stockton, and S. Hooper. 2014. "The Effects of Regulation and Litigation on a Large For-Profit Nursing Home Chain." *J. Health Polit. Policy and Law* 39 (1): 781–809.
Herd, P., and D. P. Moynihan. 2018. *Administrative Burden: Policymaking by Other Means*. New York: Russell Sage Foundation.
HHS-OIG. 2006. "Medicare Payments for Ambulance Transports." https://oig.hhs.gov/documents/evaluation/2337/OEI-05-02-00590-Complete%20Report.pdf.
Jones Day. 2021. "Supreme Court Declines to Resolve Circuit Split on Falsity Under the FCA." Jones Day, April 16.
Kaplow, L. 1992. "Rules versus Standards: An Economic Analysis." *Duke Law J.* 42 (3): 557–629.
Karpoff, J. M., D. S. Lee, and V. P. Vendrzyk. 1999. "Defense Procurement Fraud, Penalties, and Contractor Influence." *J.P.E.* 107 (4): 809–42.
Kessler, D. 2011. *Regulation versus Litigation: Perspectives from Economics and Law*. Chicago: University of Chicago Press.
Landis, J. M. 1938. *The Administrative Process*. New Haven, CT: Yale University Press.
League, R. J. 2023. "Administrative Burden and Consolidation in Health Care: Evidence from Medicare Contractor Transitions." Technical report, https://rileyleague.github.io/files/MAC_transitions.pdf.
Leder-Luis, J. Forthcoming. "Can Whistleblowers Root out Public Expenditure Fraud? Evidence from Medicare." *Rev. Econ. and Statis.*.
Lotven, A. 2022. "Advocates Seek to Delay Rsnat via Upcoming Omnibus Bill." https://insidehealthpolicy.com/daily-news/advocates-seek-delay-rsnat-upcoming-omnibus-bill.
Matsumoto, B. 2020. "Detecting Potential Overbilling in Medicare Reimbursement via Hours Worked: Comment." *A.E.R.* 110 (12): 3991–4003.
Mookherjee, D., and I. P. L. Png. 1992. "Monitoring Vis-a-Vis Investigation in Enforcement of Law." *A.E.R.* 82 (3): 556–65.
OIG (Office of Inspector General). 2021. "Fraud and Abuse Laws." https://oig.hhs.gov/compliance/physician-education/fraud-abuse-laws/.

O'Malley, A. J., T. A. Bubolz, and J. S. Skinner. 2021. "The Diffusion of Health Care Fraud: A Network Analysis." Working Paper no. 28560, NBER, Cambridge, MA.

Polinsky, A. M., and S. Shavell. 2000. "The Economic Theory of Public Enforcement of Law." *J. Econ. Literature* 38 (1): 45–76.

Sahni, N. R., B. Carrus, and D. M. Cutler. 2021. "Administrative Simplification and the Potential for Saving a Quarter-Trillion Dollars in Health Care." *J. American Medical Assoc.* 326 (17): 1677–78.

Sanghavi, P., A. B. Jena, J. P. Newhouse, and A. M. Zaslavsky. 2021. "Identifying Outlier Patterns of Inconsistent Ambulance Billing in Medicare." *Health Services Res.* 56 (2): 188–92.

SBA (Small Business Administration). 2021. "Covid-19 EIDL Program Recipients on the Department of Treasury's Do Not Pay List." Report no. 22-06, SBA, Office of Inspector General.

Shavell, S. 1984. "Liability for Harm versus Regulation of Safety." *J. Legal Studies* 13 (2): 357–74.

———. 1991. "Specific versus General Enforcement of Law." *J.P.E.* 99 (5): 1088–108.

Shepard, M., and M. Wagner. 2025. "Do Ordeals Work for Selection Markets? Evidence from Health Insurance Auto-Enrollment." *A.E.R.* 115 (3): 772–822.

Silverman, E., and J. Skinner. 2004. "Medicare Upcoding and Hospital Ownership." *J. Health Econ.* 23 (2): 369–89.

Spitalnic, P. 2018. "Certification of Medicare Prior Authorization Model for Repetitive Scheduled Non-Emergent Ambulance Transport (RSNAT)." Technical report, CMS, Baltimore.

Stigler, G. J. 1970. "The Optimum Enforcement of Laws." *J.P.E.* 78 (3): 526–36.

US GAO (Government Accountability Office). 2006. "Hurricanes Katrina and Rita Disaster Relief: Improper and Potentially Fraudulent Individual Assistance Payments Estimated to Be between $600 Million and $1.4 Billion." https://www.gao.gov/assets/gao-06-844t.pdf.

USRDS (United States Renal Data System). 2020. "USRDS 2020 Annual Data Report: Epidemiology of Kidney Disease in the United States." Bethesda, MD: NIH.

Weinstock, J., K. Purcell, K. Contreary, G. Haile, A. Goldstein, J. Coopersmith, T. Chen, and A. Asher. 2018. "First Interim Evaluation Report of the Medicare Prior Authorization Model for Repetitive Scheduled Non-Emergent Ambulance Transport (RSNAT)." Technical report, Mathematica Policy Res., Princeton, NJ.

Wilson, W. 1913. *The New Freedom: A Call for the Emancipation of the Generous Energies of a People.* New York: Doubleday.

Zeckhauser, R. 2021. "Strategic Sorting: The Role of Ordeals in Health Care." *Econ. and Philosophy* 37 (1): 64–81.