

Muehlenbachs, Lucija; Staubli, Stefan; Cohen, Mark A.

Working Paper

The Effect of Inspector Group Size and Familiarity on Enforcement and Deterrence

IZA Discussion Papers, No. 7876

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Muehlenbachs, Lucija; Staubli, Stefan; Cohen, Mark A. (2013) : The Effect of Inspector Group Size and Familiarity on Enforcement and Deterrence, IZA Discussion Papers, No. 7876, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/89925>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 7876

The Effect of Inspector Group Size and Familiarity on Enforcement and Deterrence

Lucija Muehlenbachs
Stefan Staubli
Mark A. Cohen

December 2013

The Effect of Inspector Group Size and Familiarity on Enforcement and Deterrence

Lucija Muehlenbachs

Resources for the Future, Washington DC

Stefan Staubli

*RAND Corporation,
University of Zurich and IZA*

Mark A. Cohen

*Vanderbilt University
and Resources for the Future*

Discussion Paper No. 7876

December 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Effect of Inspector Group Size and Familiarity on Enforcement and Deterrence^{*}

The paper provides new insights into the productivity of teams and the relationship between the inspector and the inspected party. Exploiting exogenous variation in the number of inspectors that are sent to offshore oil and gas platforms in the Gulf of Mexico, we find that adding an inspector does not simply result in more observed violations – it increases the severity of sanctions imposed on those violations that are detected. We also find that inspectors who are more familiar with the offender impose less severe sanctions. We only find weak evidence that increasing sanction severity deters incidents such as oil spills.

JEL Classification: Q58, K42

Keywords: inspections, enforcement, deterrence, offshore oil

Corresponding author:

Stefan Staubli
RAND Corporation
Arlington, VA 22202
USA
E-mail: sstaubli@rand.org

^{*} We thank Jessica Chu, Todd Gerarden, Madeline Gottlieb, Sharon Pailler, and Gabriel Sampson for excellent research support, Dietrich Earnhart, Catherine Hausman, Joni Hersch, Nicholas Sanders, Lance Lochner, Shanjun Li, Arvind Magesan, Devesh Rustagi, Jay Shimshack, Adam Stern, Christopher Timmins, W. Kip Viscusi, Achim Voss, and seminar participants at Vanderbilt, ETH Zurich, Duke University-North Carolina State University-RTI, and the University of Calgary for their guidance, and Tom Machado and Warren Williamson for sharing some of their institutional knowledge. A previous version of this paper was circulated as “Drill Safely Drill: The Role of Compliance Inspections on Oil & Gas Platforms in the Gulf of Mexico.” Research supported by NSF grant SES-1251916.

Following Becker’s [1968] seminal paper on the economics of crime and punishment, a burgeoning literature has developed extending and testing his model in both the street crime and regulatory context (see Polinsky and Shavell [2009] for a review). Much of the empirical economics of crime literature has focused on the incentives and behavior of *potential offenders*. Examples include the deterrent effect of increasing detection probability, such as by increasing police on the street [Corman and Mocan, 2000, Di Tella and Schargrodsky, 2004], referees on a basketball court [McCormick and Tollison, 1984] or inspections of oil transport vessels¹ and emissions at industrial facilities.² Other examples include the deterrent effect of increasing sanctions, such as imprisonment [Drago et al., 2009, Levitt, 1998] or fines [Shimshack and Ward, 2005].

A smaller, but growing, literature has focused on the incentives and behavior of the *enforcement agent*. Recent anecdotes of industry pressure on (or even “capture” of) inspectors in the nuclear, coal mining, and oil drilling industries highlight the importance of understanding the role that inspectors play in ensuring compliance. To our knowledge, all of the previous studies on incentives and behavior of enforcers are based on individual behavior and not group behavior.³ Yet, in many cases enforcement agencies send teams of enforcers, often with the express goal of minimizing corruption [Klitgaard, 1988]. This paper provides new insights into the productivity of teams and the inspector-offender relationship by examining data on inspections of offshore production facilities in the Gulf of Mexico.

The aim of the present paper is twofold. First, we study the effect that an additional inspector in an inspection team has on enforcement behavior and deterrence. Important

¹See for example, Epple and Visscher [1984], Cohen [1987], Grau and Groves [1997], and Gawande and Wheeler [1999]

²There is an extensive literature on the frequency of inspections in environmental compliance, see for example, Magat and Viscusi [1990], Laplante and Rilstone [1996], Gray and Deily [1996], Helland [1998], Earnhart [2004] or Cohen [1999] and Gray and Shimshack [2011] for reviews of this literature.

³A related literature has examined how varying the size of inspector groups has an impact on detected violations. For example, McCormick and Tollison [1984] examine the impact of increasing the number of basketball referees from two to three—and show that the accuracy of officiating improves and deters violators. Levitt [2002] found no evidence that increasing the number of NHL referees changed the number of observed infractions leading to a penalty. These studies have focused on whether groups offer more accurate assessments and their deterrent effect, and are not focused on the behavior of the enforcer and their relationship with the offender.

limitations to investigating the effect of an additional inspector are endogeneity and selection bias. Platforms that are thought to be poor performers may be sent more inspectors (in anticipation of more violations) or fewer inspectors (in anticipation that violations will be easy to detect). We use weather as an instrumental variable for the number of inspectors to identify a causal effect of group size on the level of enforcement. Inspections of offshore oil and gas platforms require the use of helicopters to take inspectors to production platforms; however, whether a helicopter can fly depends on weather conditions. When weather permits only a certain part of the Gulf to be accessed, while another part is inaccessible, then inspection plans must be modified, and in the process more inspectors are sent on fewer inspections.⁴ We find that increasing the number of inspectors in the inspection team results in an increased number of severe sanctions, implying that enforcement would be increased by adding manpower to inspections. We then use this variation in severity to examine the effect on platform safety.⁵ Using data on incidents such as oil spills, explosions, injuries and fatalities that occurred at each platform, we find no evidence that more inspectors has an effect on reducing the occurrence of reported incidents in the very short term (six and twelve months).

The second aim of this paper is to examine the effect of reducing enforcer-offender familiarity on enforcement behavior and deterrence. Using records of each inspector's inspection history, we calculate the number of times an inspector has visited the same platform before, which is a good measure for the familiarity an inspector has with a facility. We examine whether the enforcement actions and the platform safety record are affected by the number of times an inspector has visited the same platform before. However, it may be the case that more inspections of the same platform by the same inspector are driven by factors correlated with the outcome of interest. For example, poor performers have more inspections in the

⁴We first learned this through a communication with a lead inspector, but it is also borne out in the data.

⁵Apart from the broad literature on the effects of environmental enforcement on deterrence, there are other studies on the determinants of investments in safety, for example on the effect of OSHA regulations on workplace safety [Viscusi, 1986], market restructuring on nuclear safety [Hausman, Forthcoming], and profitability in the airline industry [Rose, 1990].

past and therefore are more likely to have more visits by the same inspectors. To obtain causal estimates, we exploit the closure of an inspections office that resulted in platforms in one district of the Gulf being inspected by less familiar inspectors. Since platforms in other districts were not affected by this change, we can use a difference-in-differences approach to estimate the impact of reducing familiarity on enforcement and deterrence. We find that more familiar inspectors are less likely to issue enforcement actions. Furthermore, a reduction in the familiarity between inspectors and the regulated entity results in more stringent sanctions being imposed. Yet, as in the case of inspector group size, we find no significant impact of inspector-platform familiarity on reported incidents, such as oil spills, blowouts, injuries and fatalities, within six and twelve months after an inspection. One potential explanation is that the enforcement actions available to inspectors are not stringent enough to alter investments in platform safety.

This paper relates to and builds upon several strands of literature. First, most studies on inspections and enforcement have generally focused on deterrence, i.e., whether increased enforcement activity is correlated with higher compliance and/or improved environmental performance, see Gray and Shimshack [2011] for a thorough review. This stream of literature examines deterrence at the extensive margin by measuring the effect of the frequency of inspections and less so on the effect of increasing inspection intensity, which is the focus of this paper. The investigation of the effects of increased inspection intensity is important for the design of optimal inspection policy. Second, several studies have examined various effects that might arise in a group setting. In the extreme, groupthink can result in a collective willful blindness of bad news.⁶ On the other hand, teams may be more productive than individuals [Hamilton et al., 2003] but could also be less productive in the case of free-riding [Alchian and Demsetz, 1972]. Moreover, peer pressure within a group might influence members of the team to conform to the group norm [Kandel and Lazear, 1992]. Also, Bernhard et al. [2006] and Chen and Li [2009] find that individuals are more likely to punish violators when those

⁶See [Bénabou, 2013] for a theoretical model motivated by the Columbia accident being connected to groupthink spurring overconfidence and denial.

violations occur by an “outgroup” member as opposed to an “ingroup” member. To the extent a group of inspectors identify with the regulated party, they might be less inclined to punish a violator than a group of inspectors without such a strong identification. Consistent with this hypothesis, we find that decreasing the familiarity between inspectors and the regulated entity results in more stringent sanctions being imposed.

The paper also relates to the literature on corruption and rent-seeking. Becker and Stigler [1974] noted that both criminal and regulatory offenders who face punishment for their violations have an incentive to influence the police officer or regulatory enforcement agency. Corruption among both police and regulatory enforcement agents is well documented [Rose-Ackerman, 1999]. Formal extensions to the Becker-Stigler model by Mookherjee and Png [1995] and Bowles and Garoupa [1997] have further examined trade-offs between detection probability, sanctions, and optimal enforcer compensation. While these formal models assume monetary payoffs (e.g., bribery and corruption), influence may take on more subtle forms of persuasion. A few empirical studies have examined these nonmonetary influences. Makkai and Braithwaite [1992] studied nursing home inspectors and found that those who self-identified with the industry on an attitudinal survey were more likely to issue positive compliance ratings during an inspection. Garicano et al. [2005] empirically demonstrate bias by soccer referees who are presumably under social pressure to favor home teams. In a closely related paper, Jin and Lee [2011] finds evidence of collusive behavior between inspectors and restaurants. Here, we provide a plausibly exogenous determinant of the familiarity between the enforcer and the enforced in order to identify these effects.

The paper is organized as follows. In the next section we provide background information on inspections in the Gulf of Mexico and discuss theoretical implications for optimal inspection policy. Section 2 lays out our identification strategy to examine the impact of inspector group size and familiarity on enforcement and incidents. In Section 3 we describe our data and provide some summary statistics. In Section 4 we discuss our main results. Section 5 summarizes and draws some policy conclusions.

1. BACKGROUND

1.1. *Inspections in the Gulf of Mexico*

Inspecting offshore oil and gas platforms (or also termed facilities) is a costly endeavor. In 2010, the budget for inspections of the offshore oil and gas regulator was \$23 million.⁷ However, there are over 3,000 offshore production platforms in the Gulf of Mexico alone, which must be inspected at least once every year.⁸ As of 2011, these inspections were carried out by a staff of only approximately 55 inspectors. Since the BP Deepwater Horizon spill in 2010, there have been calls for significant increases in the regulator’s budget and government enforcement activity. The 2012 fiscal year total budget for the regulator doubled to \$500 million, with the call to “hire new oil and gas inspectors.”⁹ An increase in the number of inspectors could result in an increase in the number of inspectors that are sent to each platform, but also allow for a reduction in the frequency of repeat visits of an inspector to the same platform. Both changes would increase inspection intensity and thus potentially lead to stricter enforcement actions and reduce the number of incidents. However, it may also be the case that the more times inspectors have visited the same platform, the more familiar they are with the platform’s flaws and prior violation history, making the detection of violations easier.

Inspectors perform various inspections, not only of the fixed production platforms discussed here, but also of mobile offshore drilling units (such as the Deepwater Horizon drilling rig), meters measuring oil and gas flows, and miles of pipelines. To unify the type of inspection examined in this study, we focus on the annual “production” inspections of production platforms.¹⁰ Production inspections are required for all platforms at least once annually for

⁷At the time the regulator was called the Bureau of Ocean Energy Management (BOEMRE), formerly the Minerals Management Service (MMS). On October 1, 2011, BOEMRE was split into two agencies, Bureau of Ocean Energy Management (BOEM) with the responsibility of planning and leasing, and the Bureau of Safety and Environmental Enforcement (BSEE) having responsibility for compliance inspections. Thus the data analyzed in this paper will be attributed to BSEE, however at the time of the inspections the agency was called MMS.

⁸Outer Continental Shelf Lands Act of 1953 (as amended).

⁹Budget of the United States Government, Fiscal Year 2012, p. 101.

¹⁰Note that nevertheless, we use data from all types of inspections in constructing variables on the inspec-

compliance with safety regulations, accurate recordkeeping, and proper maintenance. The primary goal of the production inspections is to verify that platforms are being operated in a safe manner according to all laws and regulations listed in the National Possible Incident of Non-Compliance (PINC) list.¹¹

Figure 1 depicts the location of the offshore oil and gas platforms across the six different districts of the Gulf (Corpus Christi, Houma, Lafayette, Lake Charles, Lake Jackson, and New Orleans). Inspector groups travel to the platforms using helicopters which are stationed at the district airports, indicated by the stars in Figure 1. An inspection group can range from one to six inspectors. A typical inspection unfolds in the following way: in approaching the platform with a helicopter, the inspector looks at the surrounding water for pollution, any drilling, or vent gas [Offshore Energy and Minerals Management, 2009]. Upon landing, the inspector walks around the platform to check its general condition, while reviewing the PINC list to determine whether violations are occurring. After the walk-around, the inspector starts the paperwork review and determines the operating ranges and safety device set points. If there is more than one inspector in the group, one inspector reviews paperwork while any other inspectors begin testing safety devices with the operator. The operator tests the devices while the inspector verifies that the safety device is “set” within the allowable range and “trips” at the set point or within the correct tolerance of the set point.

General violations and specific safety equipment failures are given “Incidents of Noncompliance” (INCs). There are three different types of INCs: (1) a warning (W), in which the operator is ordered to address the problem; (2) a component shut-in (C), requiring the operator to suspend the operation of a piece of equipment that is not functioning properly, which may or may not hinder production; and (3) a facility shut-in (F), which requires cessation of all production on the platform until the problem is mitigated and verified during a follow-up inspection. Any violations are written on the inspection form, and the INCs are written and given to the company representative, who reads the INC and is asked whether he or she has

tion history of inspectors.

¹¹An example from this list is listed in the Appendix, Section A.

any questions. The company representative signs the INC sheet and is told to return a copy within 14 days. The inspector group then returns to the district office. A different inspector in the district office verifies the data. The file clerk forwards completed inspection forms to the supervisory inspector for his or her review.

Enforcement actions taken by Bureau of Safety and Environmental Enforcement (BSEE) inspectors may be more than simple slaps on the wrist. Indeed, [Office of Inspector General, Department of the Interior, 2010a] estimates that a facility shut-in costs a large production platform over \$900,000 a day. While the bulk of enforcement actions result in the less severe INCs, warnings and component shut-ins, about 2.4% of all production inspections result in a facility shut-in—which could amount to millions of dollars annually in costs to the industry.

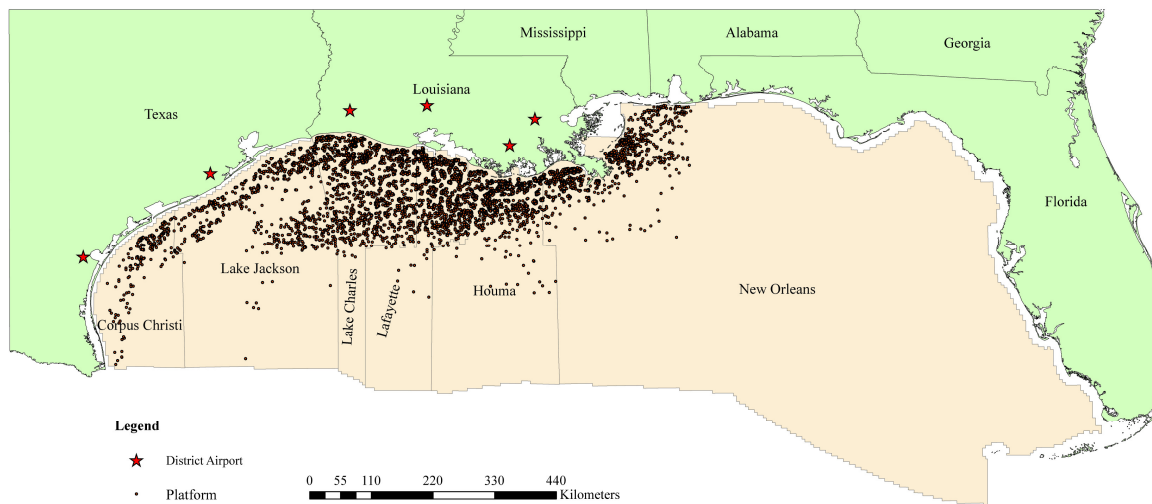


FIGURE 1.— Districts, Platforms, and District Airports in the Gulf of Mexico.

1.2. *Helicopter Flying Conditions and Closure of Corpus Christi Airport*

Since the inspection of offshore production platforms relies upon the ability of helicopters to transport inspectors to and from platforms, weather conditions may disrupt normal inspection schedules. Flights to unmanned structures are typically suspended when wind speed

is 27–30 m/s (60–67 mph), and all flights are suspended when wind speed is higher than 30 m/s (67 mph) [International Association of Oil & Gas Producers, 2005]. On average, bad weather results in inspectors only being able to fly 80% of the time [Offshore Energy and Minerals Management, 2009]. However, there are also occasions when bad weather does not completely ground flights, but instead alters the sample of platforms that can be inspected on a given day by only preventing inspections of platforms in one location while permitting inspections in another. Because the day’s inspections are decided at the inspections office, before the inspectors leave for the airport; once inspectors reach the airport, weather may preclude some of the day’s platforms from being inspected. Since the platforms to be inspected are determined before weather conditions are known, if there is bad weather restricting the flights to one area of the district, inspectors will be “doubled up” [personal communication with a lead inspector]. Therefore, weather conditions can be used as an exogenous determinant of the number of inspectors while not influencing other conditions at the platforms that are inspected.

We expect that the larger the geographical area, the more likely it is for weather to have an effect on group size. The larger the area, the more likely it is for there to be enough variation in weather that some of the planned inspections of the day would not be feasible. As shown in Figure 1, the New Orleans district covers the largest area in the Gulf of Mexico, however, drilling restrictions off of the coast of Florida have resulted in platforms not being spread across the whole district. On October 1, 2003, the Corpus Christi inspections office was closed, and all inspections for both Corpus Christi and Lake Jackson started to originate from the Lake Jackson airport. After October 2003, Lake Jackson district therefore became the district with the most dispersed platforms in the Gulf of Mexico. Thus, it is very likely that after the closure of the airport weather becomes an important determinant of inspection group size in the Lake Jackson district.

The closure of the Corpus Christi sub-district may also affect inspection intensity directly through a decline in familiarity between inspectors and the regulated entity. Specifically, by

pooling the two sub-districts, the average number of times that an inspector had inspected a platform in the past is likely to decrease. Indeed, as we will show in more detail later, we can exploit helicopter flying conditions and the closure of the Corpus Christi office to estimate the impact of inspector group size and inspector familiarity on regulatory enforcement and deterrence of incidences, such as oil spills, explosions, injuries and fatalities.

1.3. *Theoretical Implications*

There are several potential mechanisms through which increasing inspector group size and reducing inspector familiarity might increase the intensity of inspections. In the case of group size, more inspectors means that there are more eyes available to find violations and if inspections are understaffed, then with more inspectors we would expect to see an even shift increasing *all* types of enforcement actions (e.g. more warnings, component shut-ins and facility shut-ins). However group size also changes other aspects of the inspection. For example, the presence of more inspectors opens the possibility of free-riding; there could be first-order free riding in the case of inspectors not performing their duty to detect and issue INCs but there could also be second-order free riding in the case that inspectors do not punish first-order free-riding.¹² Another channel by which more inspectors would matter arises because it would be harder to “capture” the inspectors, simply because there are more inspectors to capture. If an operator protests the issuance of an INC, an inspector would have to exert more effort to issue the INC. It is the duty of the inspector not to be swayed by any protestations, so it is possible that when there are other members of an inspection group present, the inspectors might be better able to avoid operator pressure and hence more likely to write up the INC. In fact, there is anecdotal evidence that enforcement staff are subject to industry pressure. For example, after the BP Deepwater Horizon oil spill, an internal government investigation noted that inspectors in the Gulf of Mexico had been offered gifts and employment opportunities by the regulated industry [Office of Inspector

¹²See Yamagishi [1986], Heckathorn [1989], Ostrom [1990], Henrich and Boyd [2001], Boyd et al. [2003].

General, Department of the Interior, 2010b].¹³ But regulatory capture could also be driven by a more subtle form of pressure. For example, facility personnel may make comments, such as “there goes my bonus” or “my wife is sick, and I’ll lose my job,” to deter inspectors from issuing violations (Office of Inspector General, US Department of the Interior [2010]).

Similarly, the less familiar an inspector is with the operator, the less likely an inspector is to identify with the operator, and the less likely it is that the regulator is captured. However, it is also possible that inspectors who are more familiar with the platform are more likely to detect violations because they are better informed about the platform’s flaws and prior violation history. According to a post-BP Deepwater Horizon investigation by the Office of Inspector General, US Department of the Interior [2010], one inspector reported arriving at a facility where his brother, who worked for the operator elsewhere, was flown to the facility to act as the compliance officer. The inspector informed the company that he could not conduct the inspection with his brother present and another person worked with the inspector that day. In response to allegations about potential conflicts of interest, the offshore regulator issued a new ethics policy in 2011 requiring individuals in the Gulf district offices to recuse themselves in dealings with companies with a personal family or past employment relationship. According to a news report, approximately 30% of district employees recused themselves from one or more companies operating in the Gulf region in 2011.¹⁴ In this case, the less familiar an inspector is with the platform we would expect to see an increase in enforcement actions being issued.¹⁵

2. EMPIRICAL STRATEGY

2.1. *Identification Strategy for Group Size Effects*

We are interested in whether increasing the number of inspectors conducting an inspection has an effect on the enforcement actions taken as well as on the subsequent safety at the

¹³We do not have information on an inspector’s employment before or after being an inspector and so are not able to examine whether “career concerns” are driving regulatory capture.

¹⁴“Ethics Policy Quantifies Oil Industry–Gov’t Ties,” Associated Press, July 26, 2011.

¹⁵Note that our data predate these changes and end in September 2010.

platform. To examine these questions, we start by running ordinary least squares (OLS) regressions of the following type:

$$(1) \quad y_i = \varphi + \vartheta \text{Inspectors}_i + \mathbf{X}_i \boldsymbol{\varrho} + \mu_t + v_p + \varepsilon_i$$

where i denotes the inspection and Inspectors_i is the number of inspectors that conducted the inspection. We include controls for the platform, inspector, and operator in \mathbf{X}_i . Platform characteristics include age, distance to shore, water depth, production (in the month prior to the inspection), an indicator for whether the platform is a major complex, whether it is manned 24 hours a day, and the number of wells at the platform at the time of the inspection. Inspector characteristics include an indicator for whether one of the inspectors is new and the total number of inspections conducted in the past 5 years. Operator characteristics include the number of other companies that have a working interest in the lease, the percent working interest of the operator, and the number of other platforms operated by the operator, all measured at the time of the inspection. We also include operator fixed effects, v_p , to account for any non-time varying unobserved characteristics of the operator. We include an indicator for whether the inspection was unannounced, and whether the inspection occurred shortly after a hurricane. We include a year-month fixed effect, μ_t to control for district-wide changes over time. The term ε_i is a mean-zero error term, clustered at the platform operator level.

Our parameter of interest is ϑ , which measures the impact of an additional inspector on the outcome variable y_i . When examining the impact of inspector group size on enforcement, y_i denotes the aggregate number of INCs inspectors write during an inspection. However, inspectors not only might change the aggregate number of INCs they write, but they also might change the distribution of citations given; for example, they may shift from warnings to component shut-ins or from component to facility shut-ins. To investigate changes in the distribution of INCs, we break the number of INCs into subcategories including the number of warnings, the number of component shut-ins, and the number of facility shut-ins issued

during an inspection. We also examine a weighted combination of the INCs using a formula that the offshore regulator uses to calculate an Operator Safety Index: one facility shut-in is equivalent to four warnings or two component shut-ins [Slitor, 2000]. When examining the impact of inspector group size on deterrence, the outcome variable, y_i , is the count of the number of incidents such as oil spills, blowouts, injuries and fatalities that were reported on platform p within a set time period after the inspection ($t + m$ where m is six or twelve months).

A problem arises in estimating equation (1) because the number of inspectors in a group might not be randomly assigned. For example, it is possible that more inspectors are sent to facilities that are prone to violations, and therefore the covariance between the number of inspectors and ε_{ipt} would be positive. In this example, the OLS estimate of the effect of an additional inspector would be biased upwards. Therefore, to deal with the endogeneity of the number of inspectors, we instrument for the number of inspectors using the standard deviation in wind speed across the district on the day of the inspection (SDWIND_i). As can be seen in Figure 2, after the closure of the Corpus Christi office, the number of inspectors in Corpus Christi and Lack Jackson per inspection is more closely correlated with the standard deviation in wind speed across the district (the mechanism is described in Section 1.2). Figure 2 shows quarterly averages of the standard deviation in wind speed at 8 a.m. as well as the daily mean number of inspectors per inspection. As the standard deviation in wind speed across the area increases, so does the number of inspectors sent on an inspection; this is particularly pronounced after inspectors had to cover a larger geographical area.

Because larger inspection groups are sent when SDWIND is high, the difference in the resulting INCs can be used to measure the local average treatment effect (LATE) of sending more inspectors on the basis of bad weather [Imbens and Angrist, 1994]. Note that the LATE is not necessarily equal to the average treatment effect. The LATE represents a weighted average response of those platforms that had a change in inspector group size because of weather, where the weight is proportional to the number of platforms that get an additional

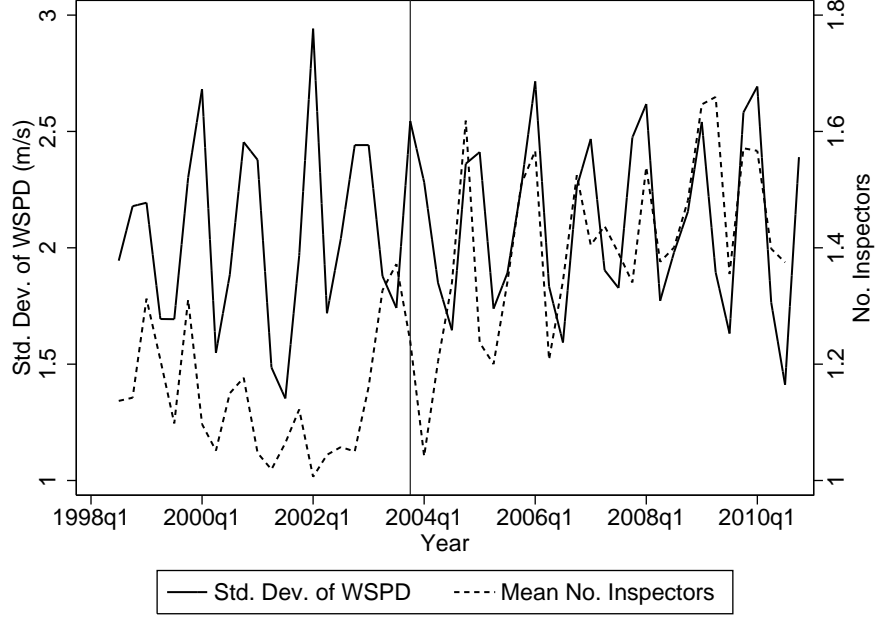


FIGURE 2.— Instrumenting for the Number of Inspectors with Variation in Wind Speed. Quarterly data obtained by averaging the 8 a.m. standard deviation of wind speed across Corpus Christi and Lake Jackson. Vertical line indicates when Corpus Christi was incorporated into the Lake Jackson district.

inspector [Imbens and Angrist, 1994, Angrist and Imbens, 1995]. The LATE does not tell us about the effect of an extra inspector on those platforms that would not get an extra inspector even in bad weather. Our first-stage equation models the relationship between wind speed and the number of inspectors:

$$(2) \quad \text{Inspectors}_i = \alpha + \beta \text{SDWIND}_i + \mathbf{X}_i \gamma + \mu_t + v_p + \varepsilon_i,$$

where SDWIND_i is the standard deviation in wind speed at 8:00 a.m. on the day the inspection takes place. Measuring the LATE requires a monotonicity assumption that variation in wind speed would not result in fewer inspectors being sent to a platform. This assumption is likely to hold in our case: if SDWIND is low, there are either no inspections because the weather is bad everywhere or fewer inspectors per inspection because the weather is good in all areas. If SDWIND is high, there are more inspectors per inspection because the weather

is good in one area and bad in another. A second assumption for the LATE is that the instrument is as good as randomly assigned and thus independent of treatment status. This assumption would be violated if platforms that are inspected when variation in wind speed is high systematically differ from the platforms that are inspected when variation in wind speed is low. In the empirical analysis we will shed light on this concern by investigating the effect of variation in wind speed on characteristics of inspected platforms. A final assumption for measuring the LATE is that variation in wind speed affects enforcement actions only through inspector group size (exclusion restriction). There would be a problem, for example, if variation in wind speed also meant that there were high winds at the platforms being inspected, which also resulted in a higher number of violations than otherwise would have occurred under normal wind speeds. This is unlikely to be the case because if weather were bad enough to cause structural damage, helicopters would be grounded and we would have no observations of inspections in the first place. Nonetheless, in the empirical analysis we will find that different measures of weather (mean wind speed across the district on the day of the inspection, wind speed closest to the platform, and mean wave height across the district) do not determine the number of enforcement actions. Nonetheless, we also include an indicator for inspections that occurred after a hurricane in equation (2).

2.2. Identification Strategy for Familiarity

To examine whether enforcement actions depend on the inspector’s familiarity with the inspected party, we use data on each inspector’s inspection history. Based on these data we can calculate the number of times an inspector has visited the same platform in the past five years, which captures the familiarity an inspector has with that particular platform. To test whether inspector familiarity has an impact on enforcement actions, we could simply regress the number and type of enforcement actions an inspector gives during an inspection on our measure of familiarity. However, the estimates from this regression are likely to be biased for two reasons. First, poor performers or older platforms would have had more inspections in the

past and might also have had more enforcement actions. This would bias estimates in favor of a finding that more inspections by the same inspector result in more enforcement actions. Second, if inspectors have any leeway in choosing where to inspect, the repeat assignment of inspectors to the same platform may be driven by unobservables correlated with the outcome we are interested in.

To address these concerns, we exploit exogenous variation in the number of times an inspector has visited the same platform created by the closure of the Corpus Christi sub-district office. With the closure of the airport, Corpus Christi was integrated into Lake Jackson and the number of platforms inspected from Lake Jackson airport increased significantly. As a consequence, the average number of times that an inspector had inspected a platform in the past decreased in Lake Jackson after the airport closure. We rely on a difference-in-differences (DID) approach to investigate how the decreased number of prior visits by the same inspector affected enforcement actions taken and subsequent safety at the platform. The first difference is over time, as the number of prior visits by the same inspector decreased after the airport closure. The second difference is across districts; only platforms in Lake Jackson and Corpus Christi were affected by the airport closure. These platforms define the treatment group. The platforms in all other regions define the control group.

The DID comparison is implemented by estimation regressions of the following type:

$$(3) \quad y_i = \alpha + \beta \text{LJCC}_p \times \text{Post}_t + \gamma \text{LJCC}_p + \mathbf{X}_i \delta + \mu_t + v_p + \varepsilon_i,$$

where LJCC_p is a dummy equal to one for platforms in Lake Jackson (or the Corpus Christi sub-district) and zero for platforms elsewhere in the Gulf and Post_t is a dummy equal to one after the airport closure, zero otherwise. Note that with year-month fixed effects, μ_t , we do not include the Post indicator by itself. The outcome variable y_i either denotes the average number of repeated visits of the inspector team in the past five years, the number of INCs by the subcategories inspectors issue during an inspection, or the number of incidents

reported within a given time period after the inspection. Our parameter of interest is β , which measures the effect of the airport closure on treated platforms relative to control platforms using variation over time. We estimate equation 3 using data from January 2000 to September 2003 as the pre-period and October 2003 to December 2006 as the post period.

The identifying assumption is that trends in y_i would have been identical in treatment and control platforms in the absence of the airport closure. To shed light on the validity of this assumption, we generalize equation (3) by replacing $LJCC_p \times Post_t$ with a set of treatment times year interaction terms using all years of available data (1998 to 2010). Each coefficient of these interaction terms measures the difference in the outcome variable in the Corpus Christi and Lake Jackson district relative to all other districts in a given year relative to the reference year (1998). The year 2003 is a transition year since the airport was closed in October 2003. The interaction terms for the years 1999-2002 should be zero if the parallel trend assumption holds. The interaction terms for the years 2004-2010 measure how the impact of the airport closure evolves over time.

3. DATA AND SAMPLE CONSTRUCTION

3.1. *Data*

For our analysis we combine data from several sources. Records of inspections were provided by the Bureau of Safety and Environmental Enforcement (BSEE). The dataset contains information on every inspection conducted in the Gulf of Mexico and includes information on the count and type of enforcement actions (INCs) taken. The dataset also contains the full names of the inspectors conducting the inspection. This information is used to calculate the number of times an inspector visited the platform in the past five years prior to the inspection, which is our measure for familiarity between the inspector and the facility. We also calculate the number of enforcement actions that a platform received in the past. We construct these variables using the full dataset dating back to 1986. BSEE also provided us with a dataset containing all serious accidents, fatalities, injuries, explosions, and fires. Incidents include fires, explosions, blowouts, loss of well control underground and at the surface, collisions,

spills of oil or drilling fluids, equipment failure, injuries, and fatalities. In the analysis we examine all incidents together, and then injuries and fatalities.¹⁶ The regulations for incident reporting were made more stringent on July 17, 2006, when operators were required to report not only serious incidents, but also incidents that had the potential to be serious (e.g., any incident involving structural damage to a platform, injury that led to an evacuation or days away from work, or property damage exceeding \$25,000).¹⁷

Two datasets¹⁸ containing platform characteristics from the Bureau of Ocean Energy Management (BOEM) were used to create a panel of every platform in the Gulf of Mexico from the year the platform was first installed through 2010. The dataset contains information on the distance to shore, water depth, lease number, and whether or not there are personnel on board 24 hours a day or whether the platform is considered to be a “major complex”¹⁹ Production data for all wells in the Gulf of Mexico from 1996 to 2010 were also obtained online from BOEM.²⁰ This dataset contains a unique well identifier (the API well number), monthly volume of oil and gas produced, and days on production. Using another dataset available from BOEM on well characteristics,²¹ we assigned a platform identifier to the API well numbers so that we could aggregate the monthly well production to monthly platform production. Using the date that the wells were drilled we created a variable for the number of wells at each platform at the time of the inspection.

Data on the ownership of the lease and the designated operator of a lease were obtained online from BOEM.²² These data contain the working interest of all owners of Gulf leases, including all ownership changes from the assignment date of the lease to present. We con-

¹⁶If an inspector discovers that an accident occurred and was not filed within 15 days, the INC they are required to issue is a warning, but nonetheless, incidents are self-reported by platforms and there is potentially under reporting.

¹⁷30 CFR Part 250, Final Rule (FR 19640), Minerals Management Service, U.S. Department of the Interior.

¹⁸“Platform Masters” and “Platform Structures” found online at https://www.data.boem.gov/homepg/data_center/platform/platform.asp

¹⁹Major complex is defined as a platform that has at least one structure with at least six completions or two pieces of production equipment

²⁰Production Data: http://www.data.boem.gov/homepg/pubinfo/freecsci/product/freeprod_ogora.asp

²¹Borehole Data: https://www.data.boem.gov/homepg/data_center/well/well.asp

²²Lease Data: https://www.data.boem.gov/homepg/data_center/leasing/leasing.asp

structed variables for the number of companies that had an interest in the lease in that year, and the mean percentage working interest in the lease. At any one point in time, however, there is only one designated operator of the lease. The operator, as defined by BOEM, is either the leaseholder or the party designated (and approved) to operate a portion of a given lease.²³ A record of platform operators (received from BOEM) was used to determine the operator of a platform at the time of the inspection. BOEM gives each subsidiary of a company a different company identification number. For example, Shell Offshore Inc. has 10 subsidiaries in the Gulf of Mexico (including, for example, Shell Consolidated Energy Resources Inc., Shell Deepwater Development Inc., Shell Oil Company), each of which has a unique company number. In order to match the subsidiaries to their parent companies, we used an unofficial list of parent-subsidiaries obtained from BOEM. For the remainder of unmatched observations, the parent company found in BOEM’s operator safety summaries was used. With this information we constructed variables for the number of platforms that the parent company operates, and the working interest of the platform operator.

Finally, wind speed data over the Gulf were obtained from the National Data Buoy Center (NDBC) of the National Oceanic and Atmospheric Administration (NOAA). Buoy stations record hourly standard meteorological and wave data. Wind speed is the most often recorded variable whereas temperature and wave height are often missing.²⁴ There are eight NDBC deployed buoys stationed off the coast of Texas. Only five of these buoys have data spanning the dates studied. We use the wind speed recorded at 8:00 a.m. of the day of the inspection because helicopters transporting inspectors attempt to take off earlier than 7:15 a.m. [Offshore Energy and Minerals Management, 2009].²⁵

²³A platform operator is typically the responsible party in the event of an oil spill. However, in some instances a platform ties in production from a subsea lease that is miles away and could be leased to a different operator. Under U.S. government regulations, all three parties (the surface platform operator, the subsea lessee, and the pipeline right-of-way holder) are required to show oil spill financial responsibility.

²⁴Although visibility, cloud ceiling, and the presence of lightning also determine whether flying is an option, it is not possible to obtain these data over the Gulf of Mexico, and our estimation strategy depends on distinguishing weather across the district.

²⁵We convert wind speeds to 20 meters above the surface using the formula $WSPD(20 \text{ meters}) = WSPD(z)(\frac{20}{z})^{1/7}$ where z represents the height above the surface at which observations are collected [U.S. Army Coastal Engineering Research Center, 1984].

3.2. Sample Construction

We construct different subsamples for each of the two estimation strategies. For our instrumental variable strategy we use production inspections of offshore platforms in Lake Jackson after it subsumed the Corpus Christi subdistrict in October 2003 until December 2010. We chose Lake Jackson because since 2003, compared to other districts in the Gulf of Mexico, the platforms in Lake Jackson are spread over the largest area (see Figure 1 for a map of the Gulf). In the other districts, which have more platforms and cover smaller areas, bad weather would be more likely to result in the grounding of all flights, not just the grounding of flights to a portion of the district. Therefore, in Lake Jackson it is more likely for there to be variation in weather across the district, and thus more likely for weather to affect the number of inspectors sent to a platform than in a geographically smaller district. Our sample for the DID estimation strategy consists of production inspections across all districts in the Gulf between January 1998 and December 2010. However, the analysis focuses primarily on production inspections between January 2000 and December 2006, given that the airport closure took place in October 2003.

We focus on annual production inspections, which account for 81% of the 13 different types of facility inspections. We restrict the sample to only include these inspections because they are more uniform in how they are conducted and in the type of INC issued. Nevertheless, to ensure we construct the variables indicating a particular inspector’s history correctly, we include data on all inspection types. Also, when the inspection has more than one inspector, the inspector variables are averaged over all inspectors in the group. For example, the number of prior inspections in the past five years ($\text{Past Inspections}_I$) is the number of prior inspections averaged over all inspectors in the group. These variables are cumulative including all inspections from 1986 to the current inspection.

Table I reports summary statistics for the instrumental variable sample and the DID sample by treatment and control region, before (January 2000 to September 2003) and after the airport closure (October 2003 to December 2006). The variables for wind speed are

only collected for the instrumental variable sample. As the first row shows, the mean of our instrument for group size, SDWIND is 1.89 m/s. However, there is substantial variation across time: the standard deviation of SDWIND is 0.923 and the minimum and maximum values are 0.269 m/s and 6.29 m/s, respectively (not reported). Because information on weather is missing for some inspections, the instrumental variable sample is smaller than the DID sample for the Lake Jackson/Corpus Christi district after the airport closure (column 3). Despite the different sample sizes, there are only small differences in terms of observable characteristics between the IV sample and the DID treatment, post-period sample.

A comparison of inspection and inspector characteristics across the different columns of the DID sample reveals that before the airport closure there are only minor differences in the number of INCs, the severity of the INCs, and the number of incidents and injuries and fatalities between Lake Jackson/Corpus Christ and the rest of the Gulf. The pre-reform differences in inspection outcomes are likely to be driven in part by differences in platform characteristics. In particular, platforms in treatment districts tend to be further away from the shore, have greater water depth, and are more likely to be active. After the airport closure INCs are generally lower in treatment and control districts, except for component shut-ins in treatment districts which are higher compared to the pre-period. The number of incidents and injuries and fatalities are increasing over time, which is partly driven by the 2006 change which increased the number of incidents that were required to be reported (as described in section 3).

4. RESULTS

4.1. *Group Size Effects on Enforcement and Deterrence*

Looking simply at the mean number of INCs per inspection by the number of inspectors conducting the inspection (Table II), there is a positive correlation between the number of inspectors and the severity of the enforcement actions taken. For example, the probability of a component shut-in increases from 0.418 to 0.527 when moving from one to two inspectors, and increases further to 1.124 with three inspectors and 1.214 with four or more. Of course,

TABLE I
SUMMARY STATISTICS

	IV	DID			
		Treatment		Control	
		Pre	Post	Pre	Post
	(1)	(2)	(3)	(4)	(5)
Outcome Variables:					
No. Warnings	.455	.547	.403	.43	.349
No. Component Shut-ins	.483	.381	.461	.419	.293
No. Facility Shut-ins	.0339	.0336	.0165	.0277	.0191
Weighted INCs	1.56	1.44	1.39	1.38	1.01
Incidents 12 Mo. After	.109	.0408	.0624	.0441	.105
Injuries & Fatalities 12 Mo. After	.0115	.0129	.0182	.0085	.0087
Weather Variables:					
SDWSPD (m/s)	1.89				
Mean WSPD (m/s)	7.27				
WSPD at Closest Buoy	7.32				
Hurricane (2 months post)	.121	.0272	.173	.028	.164
Unannounced (Indicator)	.0015	.0637	.0035	.0057	.003
No. Inspectors	1.36	1.12	1.28	1.48	1.58
Time Since Last Insp. (days)	342	333	345	337	340
Inspector Characteristics:					
New Inspector (Indicator)	.0314	.0329	.0416	.0159	.0237
Past Inspections _I (100s)	35.2	35.3	35	35.7	35.6
Cum. Inspections of P by I, past 5 yrs	.598	1.39	.528	.614	.689
Platform Characteristics:					
Platform Age (years)	14.9	12.8	13.9	19.4	20.8
Distance to Shore (miles)	43.8	44.2	43	31.5	31.8
Water Depth (1,000 feet)	.191	.148	.169	.112	.117
Production (mmBOE in month prior)	.0357	.0362	.0407	.0378	.0273
Major Complex (Indicator)	.669	.598	.644	.447	.438
Manned 24 hrs (Indicator)	.248	.195	.212	.18	.181
Number of Wells (t)	4.1	3.85	3.96	4.78	4.49
Inactive (Indicator)	.26	.234	.238	.319	.37
Ownership and Operator:					
No. of Lessees (t)	2.58	2.37	2.46	2.16	2.09
Working Interest of O (%) (t)	57	50.3	54.1	59.5	61.6
No. of P Operated by O (t)	127	135	133	213	188
Observations	2,008	1,398	1,153	10,499	8,639

Notes: IV-sample includes production inspections on offshore facilities in Lake Jackson/Corpus Christi October 2003 to August 2010. DID-sample includes production inspections in Lake Jackson/Corpus Christi (treatment group) and other districts of the Gulf (control group) between January 2000 to September 2003 (pre-period) and between October 2003 to December 2006 (post-period). P=Platform, O=Operator, I=Inspector, (O,I)=Inspections of Operator O by Inspector I. For all inspector variables (I), if the inspection was conducted by a group, the variable is the average over all group members.

these numbers do not capture the causal impact of group size on enforcement but may reflect that more inspectors are sent to platforms that are thought to be poor performers. Consistent with this view, Table II shows that there is a positive association between the number of inspectors and overall number of incidents and injuries or fatalities occurring. For example, the probability of any incident occurring 12 months after an inspection increases from 0.094 to 0.126 when a second inspector is added and increases further to 0.213 with three inspectors and 0.286 with four or more. We therefore rely on variation in weather across the Lake Jackson/Corpus Christi district after the closure of the airport, which provides us

with variation in the number of inspectors that is independent of unobservable platform characteristics.

TABLE II
MEANS (AND STANDARD DEVIATIONS) BY NUMBER OF INSPECTORS

	One (1)	Two (2)	Three (3)	Four+ (4)
No. Warnings	.437 (.993)	.448 (1.061)	.742 (1.418)	.714 (1.139)
No. Component Shut-ins	.418 (1.238)	.527 (1.368)	1.124 (2.421)	1.214 (1.477)
No. Facility Shut-ins	.021 (.211)	.049 (.250)	.146 (.490)	.071 (.267)
Weighted INCs	1.357 (3.245)	1.697 (3.746)	3.573 (6.250)	3.429 (4.603)
Incidents 12 Mo. After	.094 (.368)	.126 (.486)	.213 (.593)	.286 (.469)
Injuries & Fatalities 12 Mo. After	.008 (.103)	.020 (.234)	.022 (.149)	.000 (.000)
Observations	1396	509	89	14

Notes: Sample includes production inspections in Lake Jackson October 2003 to August 2010. Means are reported with standard deviations in parentheses. There is only one observation with six inspectors, and none with five.

Table III reports the first stage relationship between inspectors and the standard deviation in wind speed in the instrumental variables estimation. The specification in column 1 only includes year-month fixed effects, the specification in column 2 also adds operator fixed effects, and the specification in column 3 adds all exogenous variables described in detail in section 2.1. For our purposes, the important finding is that the number of inspectors is positively related to wind speed variability, consistent with Figure 2. Independent of the specification, we find a strong first stage relationship; parameter estimates are statistically significant and remarkably similar whether using additional covariates or not. All specifications also pass the weak instrument test given that the F-statistics is larger than the threshold value of 10 [Stock and Staiger, 1997]. As a placebo test, we use the standard deviation in wind speed exactly one year after the inspection as the instrumental variable (column 4). In this case, the coefficient on the standard deviation is insignificant in the first stage regression, the F-stat for weak instrument is 1.23, and all coefficients in the second stage are insignificant. Using the standard deviation in wind speed the day before (column 5) *and* the day after the inspection (column 6) results in a first stage F-statistic of 0.14 and 0.26, respectively, and all affected coefficients are insignificant.

Table IV reports on our main regression results relating the number of inspectors to the resulting enforcement actions. This table shows both OLS and IV results when including the full set of control variables, but the results are very similar for specifications that only include year-month fixed effects or year-month and operator fixed effects. The OLS estimates indicate that adding an additional inspector has no significant effect on the number of warnings issued, but leads to a significant increase in component shut-ins, facility shut-ins, and the weighted combination of enforcement actions. The IV estimates are larger in magnitude than the OLS estimates, but they are also less precisely estimated. As a consequence, in the IV estimation only the most stringent enforcement actions, facility shut-ins, are statistically significantly more likely to occur with more inspectors. The point estimate suggests that adding an inspector increases the number of facility shut-ins by .243.

The coefficient on the number of inspectors in Table IV is always larger when using the IV specification compared to OLS. Besides endogeneity, there could be two other reasons driving this result—measurement error in group size and non-linearity in the effect of group size. Our data contain detailed records of the number of inspectors on the platform and therefore measurement error in the number of inspectors is not as likely as non-linearity in the effects. The IV estimates use variation in the inspection group size induced by wind speed, whereas OLS estimates use all the variation. If weather affects a particular level shift in the group size and this level shift has a larger effect on enforcement, then the IV and the OLS estimates will differ. Appendix B examines this issue in more detail and provides evidence that non-linearities can explain the difference between IV and OLS estimates.

Table V reports the estimates on the effect of inspector group size on reported incidents such as oil spills, injuries and fatalities occurring within 6 and 12 months of any inspection. The first four columns report on all incidents, while the remaining columns report on a subset—only those involving injuries or deaths. As before, the table shows both OLS and IV results when including the full set of control variables. The OLS estimates are in all cases very close to zero and statistically insignificant. The IV estimates are negative in

TABLE III
FIRST STAGE: NUMBER OF INSPECTORS

	Main estimates			Placebo tests		
	(1)	(2)	(3)	(4)	(5)	(6)
Std. Dev. WSPD	.061** (.024)	.057** (.026)	.064*** (.023)	-.016 (.015)	.006 (.012)	-.008 (.018)
Year-Month Effects	Yes	Yes	Yes	Yes	Yes	Yes
Operator Effects	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes	Yes
n	2,008	2,008	2,008	2,008	2,008	2,008
Mean of Dep. Var.	1.364	1.364	1.364	1.364	1.364	1.364
R ²	.114	.174	.248	.242	.242	.242
First Stage F-stat.	13.19	11.07	15.1	1.23	.14	.26

Notes: Sample includes production inspections in Lake Jackson between October 2003 and September 2010. Columns represent OLS coefficients. Dependent variable is the number of inspectors. Robust standard errors are clustered by operator (for 91 operators). Controls include platform's age, distance to shore, water depth, previous month's production, count of wells, operator's number of leases, platforms and working interest, inspector's count of past inspections, indicators for whether the platform is inactive, a major complex, manned-24 hours, was in Corpus Christi, and whether the inspection was unannounced, conducted by a new inspector, or occurred after a Hurricane. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

TABLE IV
IMPACT OF GROUP SIZE ON ENFORCEMENT

	Warnings		Component		Facility		Weighted	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
No. Inspectors	.069 (.054)	.154 (.723)	.216*** (.053)	.517 (.610)	.036** (.015)	.243** (.111)	.644*** (.153)	2.159 (1.720)
Year-Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Operator Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	2,008	2,008	2,008	2,008	2,008	2,008	2,008	2,008
R ²	.190		.209		.106		.214	
Mean of Dep. Var.	.455	.455	.483	.483	.034	.034	1.556	1.556

Notes: Sample includes production inspections in Lake Jackson between October 2003 and September 2010. Dependent variables are the number of warnings (W), component shut-ins (C), facility shut-ins (F), or *Weighted*, a weighted combination of all INCs (Weighted=W+2C+4F). In the case of the IV estimates, *No. Inspectors* is instrumented for using weather. Robust standard errors are clustered by operator (for 91 operators). Controls include platform's age, distance to shore, water depth, previous month's production, count of wells, operator's number of leases, platforms and working interest, inspector's count of past inspections, indicators for whether the platform is inactive, a major complex, manned-24 hours, was in Corpus Christi, and whether the inspection was unannounced, conducted by a new inspector, or occurred after a Hurricane. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

all specifications and quite sizeable in magnitude, but the coefficients are not statistically significant. Together, the findings from tables IV and V suggest that while adding inspectors to an inspection leads to more stringent enforcement actions, increasing inspector group size seems not to have an impact on platform safety following an inspection. We explore possible explanations for this finding in the concluding section.

TABLE V
IMPACT OF INSPECTOR GROUP SIZE ON DETERRENCE

	All Incidents				Injuries & Fatalities			
	6 Mo. After		12 Mo. After		6 Mo. After		12 Mo. After	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
No. Inspectors	.005 (.012)	-.135 (.114)	.000 (.020)	-.112 (.190)	-.001 (.004)	-.083 (.084)	-.002 (.007)	-.013 (.054)
Year-Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Operator Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	2,008	2,008	2,008	2,008	2,008	2,008	2,008	2,008
R ²	.172		.241		.232		.322	
Mean of Dep. Var.	.055	.055	.109	.109	.006	.006	.011	.011

Notes: Sample includes production inspections in Lake Jackson between October 2003 and September 2010. Dependent variables are the number of all incidents and a subset of incidents – injuries & fatalities – within 6 months and 12 months after an inspection. In the case of the IV estimates, *No. Inspectors* is instrumented for using weather. Robust standard errors are clustered by operator (for 91 operators). Controls include platform’s age, distance to shore, water depth, previous month’s production, count of wells, operator’s number of leases, platforms and working interest, inspector’s count of past inspections, indicators for whether the platform is inactive, a major complex, manned-24 hours, was in Corpus Christi, and whether the inspection was unannounced, conducted by a new inspector, or occurred after a Hurricane. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

One potential concern with estimating the model is that platforms that are inspected when variation in wind speed is high systematically differ from platforms that are inspected during good weather. Such selection behavior would violate the assumption that the instrument is as good as randomly assigned. As shown in Table VI, baseline characteristics, such as the distance to the shore or the age of the platform, do not to vary across platforms inspected on highly variable days, suggesting that this concern is unfounded. It is also not the case that the instrument is affecting the familiarity of the inspection group and the inspected platform (last column of Table VI) implying that our estimate of the effect of group size on enforcement is not confounded by a change in familiarity. Another potential worry is that wind speed may have a direct effect on the outcome of an inspection, which would violate the exclusion restriction. We show in Table VII that the mean wind speed in the district on the day of the inspection, the wind speed closest to the platform, and the average wave height do

TABLE VI
BASELINE CHARACTERISTICS BY STANDARD DEVIATION IN WIND SPEED

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dist.	Age	Depth	Prod.	Major	Manned	Wells	INCs(t-1)	C.C.	Repeat
Std. Dev. WSPD	.553 (.731)	-.056 (.276)	.008 (.015)	.003 (.007)	.001 (.013)	-.019 (.014)	-.214 (.140)	.056 (.072)	.001 (.009)	-.004 (.024)
Year-Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Operator Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	2,008	2,008	2,008	2,008	2,008	2,008	2,008	2,008	2,008	2,008
R ²	.385	.410	.377	.300	.206	.273	.349	.113	.426	.261

Notes: Columns represent OLS coefficients with robust standard errors clustered by operator (for 91 operators). Dependent variables are platform characteristics: distance to shore, age, depth, past month's production, whether the complex is a major complex, whether it is manned 24 hours a day, the number of wells present, a weighted average of the INCs from the most recent prior inspection (INCs=W+2C+4F), an indicator for whether the platform was in the Corpus Christi sub-district (CC), and the number of repeat visits of an inspector to the same platform in the past 5 years. Year-month and operator fixed effects included in all regressions (coefficients not reported). Sample includes production inspections in Lake Jackson October 2003 to August 2010. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

not have an effect on the enforcement action taken.²⁶ This also follows the reasoning that if weather were bad enough to cause structural damage, helicopters would be grounded and we would have no observations of inspections in the first place. It still could be possible that high variation in wind speed coincides with incidents of non-compliance issued because of damage from hurricanes. We find that an indicator for inspections that occurred after a hurricane are statistically insignificant in determining the number of inspectors or the enforcement action taken.²⁷

4.2. *Effect of Familiarity on Enforcement and Deterrence*

To examine the impact of familiarity on enforcement and deterrence, we exploit the closure of the Corpus Christi airport in October 2003 after which all inspections for both Corpus Christi and Lake Jackson started to originate from the Lake Jackson airport. As the left Panel of Figure 3 illustrates, there is a sizeable drop in repeat visits of an inspector to the same platform in the Corpus Christi and Lack Jackson district after the airport closure, which suggests that there was also a drop in the familiarity of inspectors and the inspected

²⁶In the specification shown in Table VII, all covariates that are included in the main regression (Table IV) are also included; however, not including these covariates also does not result in statistical significance on mean wind speed on the day of the inspection, the wind speed measured at the buoy closest to the platform, nor the average wave height. Due to space constraints the weighted average of the INCs is not included, but all coefficients are also statistically insignificant in the weighted average INC regression.

²⁷We also tried various intervals ranging from one week to 16 weeks after a hurricane. The final specifications presented in the paper include an indicator for any inspection that occurred within eight weeks after Hurricanes Lili, Ivan, Dennis, Katrina, Rita, Gustav, and Ike.

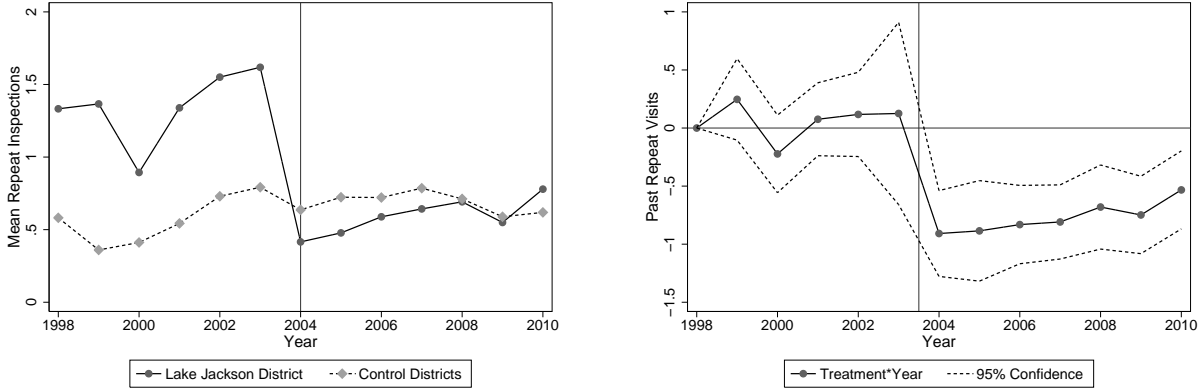
TABLE VII
DIRECT EFFECT OF WEATHER ON ENFORCEMENT

	Warnings			Component			Facility		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mean WSPD	.003 (.014)			.006 (.013)			-.002 (.003)		
WSPD at Closest Buoy		.002 (.009)			.006 (.012)			.000 (.003)	
Mean Wave Height			-.020 (.065)			-.017 (.061)			-.006 (.017)
Year-Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Operator Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	2,008	2,000	1,787	2,008	2,000	1,787	2,008	2,000	1,787
R ²	.188	.189	.208	.202	.202	.205	.101	.101	.103
Mean of Dep. Var.	.455	.456	.464	.483	.484	.485	.034	.034	.035

Notes: Sample includes production inspections in Lake Jackson between October 2003 and September 2010. Dependent variables are the number of warnings (W), component shut-ins (C), and facility shut-ins (F). Robust standard errors are clustered by operator (for 91 operators). Controls include platform's age, distance to shore, water depth, previous month's production, count of wells, operator's number of leases, platforms and working interest, inspector's count of past inspections, indicators for whether the platform is inactive, a major complex, manned-24 hours, was in Corpus Christi, and whether the inspection was unannounced, conducted by a new inspector, or occurred after a Hurricane. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

platforms. To examine the effect of the airport closure in more detail, we replace $LJCC_p \times Post_t$ in equation 3 with a set of $LJCC_p$ -year interaction terms. The right panel of Figure 3 plots the coefficient estimates of these interaction terms. The estimates for the period before the airport closure are not significantly different from zero, providing evidence for the parallel trend assumption. Consistent with the left panel, there is a significant and long-lasting decline in repeat visits after 2003; even seven years after the airport closure repeat visits are significantly lower in the Corpus Christi and Lake Jackson district.

How did the reduction in inspector familiarity with the inspected entity affect enforcement actions? The answer to this question is provided in Table VIII which reports estimates of equation (3) for different enforcement actions (as well as repeat visits and inspector group size) using data from the years 2000-2006, three years before and after the airport closure. Column 1 shows a statistically significant reduction in the number of repeat visits in Lake Jackson-Corpus Christ in the post-closure period. On the other hand, the average inspector group size is not affected by the airport closure, as illustrated in column 2. Columns 3-5 suggest that the reduction in repeat visits, and not group size, led to a significant increase in component shut-ins and in the weighted average of INCs.



(a) Mean number of visits by inspector prior to inspection.

(b) Difference-in-Difference Coefficients for (Treatment * Year Indicators)

FIGURE 3.— The Corpus Christi inspections office was closed in 2003, after which the number of repeat inspections significantly decreased.

TABLE VIII
IMPACT OF REDUCING REPEAT VISITS ON ENFORCEMENT ACTIONS

	Repeat (1)	Inspectors (2)	Warnings (3)	Component (4)	Facility (5)	Weighted (6)
LJCC \times Post Closure	-.883*** (.210)	.048 (.043)	-.055 (.080)	.231*** (.077)	-.007 (.009)	.381* (.219)
LJCC	.751*** (.143)	-.338*** (.029)	-.030 (.067)	-.215*** (.045)	-.009 (.008)	-.496*** (.163)
Year-Month Effects	Yes	Yes	Yes	Yes	Yes	Yes
Operator Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
n	21,689	21,689	21,689	21,689	21,689	21,689
R ²	.175	.183	.114	.109	.037	.138
Mean of Dep. Var.	.689	1.488	.404	.368	.024	1.237

Notes: Sample includes all production inspections in the Gulf of Mexico between January 2000 and December 2006. The indicator LJCC includes platforms that were in the Lake Jackson or Corpus Christi sub-districts. Post Closure= 1 if the inspection occurred after October 2003. The dependent variables are Repeat (=the number of prior visits the inspector made to the same platform in the past 5 years); Inspectors=the number of inspectors; the number of Warnings (W) issued; Component shut-ins (C); Facility shut-ins (F); or *Weighted*, a weighted combination of all INCs (Weighted=W+2C+4F). All specifications include all control variables as discussed in the main text. Robust standard errors are clustered by operators (for 181 operators). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Overall these results lend support to the claim that the less familiar an inspector is with the operator, the less likely an inspector is to identify with the operator, and the less likely it is that the inspector is captured. To examine whether the increase in inspection intensity has an effect on platform safety, Table IX reports OLS estimates of equation 3 using as outcome variables the number of incidents (including oil spills, blowouts, and injuries and fatalities) and the number of injuries and fatalities occurring within 6 and 12 months after an inspection. Unlike the case of inspector group size, we find that reducing inspector familiarity leads to a

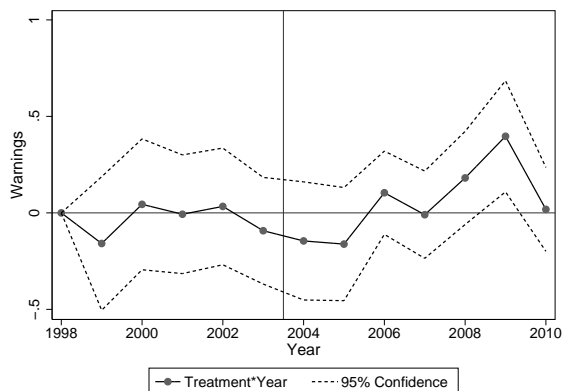
significant decrease in reported incidents 6 and 12 months after an inspection, suggesting that reducing familiarity may be a more effective measure to improve platforms' safety behavior than increasing group size. Columns (3) and (4) indicated that reducing inspector familiarity has no effect on the number of injuries and fatalities; none of the estimates of the interaction term $LJCC_p \times Post_t$ are significantly different from zero.

TABLE IX
IMPACT OF REDUCING REPEAT VISITS ON DETERRENCE

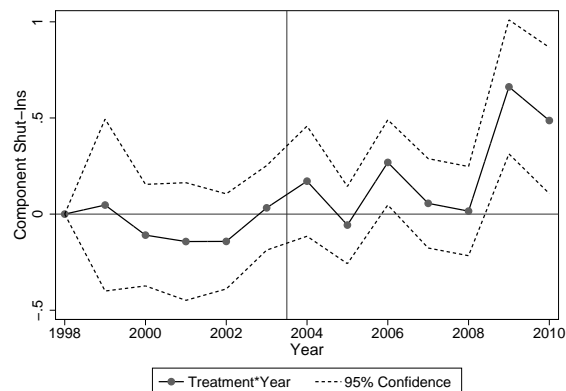
	All Incidents		Injuries & Fatalities	
	(1) 6 Mo. After	(2) 12 Mo. After	(3) 6 Mo. After	(4) 12 Mo. After
LJCC \times Post Closure	-.015** (.007)	-.023* (.012)	.004 (.006)	.002 (.011)
LJCC	-.003 (.006)	-.008 (.010)	.002 (.003)	.005 (.005)
Year-Month Effects	Yes	Yes	Yes	Yes
Operator Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
n	21,689	21,689	21,689	21,689
R ²	.083	.138	.012	.019
Mean of Dep. Var.	.028	.058	.005	.009

Notes: Sample includes all platforms in the Gulf of Mexico. The indicator LJCC includes platforms that were in the Lake Jackson or Corpus Christi sub-districts. Post Closure= 1 if the inspection occurred after October 2003. Dependent variables are the number of all incidents and a subset of incidents – injuries & fatalities – reported within 6 months and 12 months after an inspection. All specifications include control variables as discussed in the main text. Robust standard errors are clustered by operators (for 181 operators). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

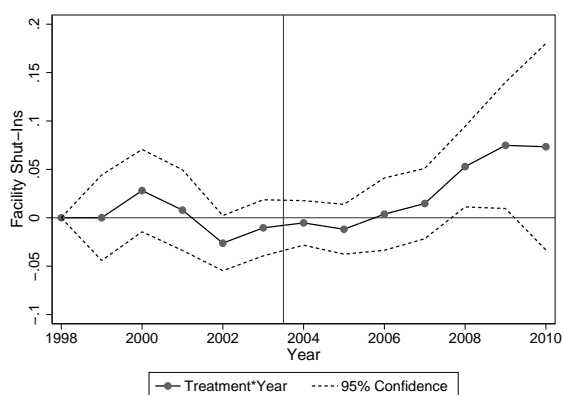
The estimates in Tables VIII and IX are only informative about the impact of reducing inspector familiarity if trends in INCs and incidents would have been similar in control and treatment districts in the absence of the airport closure. To shed light on the validity of this assumption, we replace the indicator for the post period in equation 3 with year indicators, and plot the estimates of the interaction terms from regressions of enforcement actions, incidents within 12 months, and injuries and fatalities within 12 months in Figure 4. In all six subfigures, the coefficients in the pre-period are not significantly different from zero, suggesting that trends in the outcome variables of interest are similar in the Lake Jackson-Corpus Christi district and the control districts prior to the airport closure. Consistent with the estimates in Tables VIII and IX, there are only small changes in enforcement actions right after the airport closure. Component shut-ins and weighted INCs tend to be higher after the airport closure, but the coefficient estimates are not statistically significant except for the last two years (2009 and 2010).



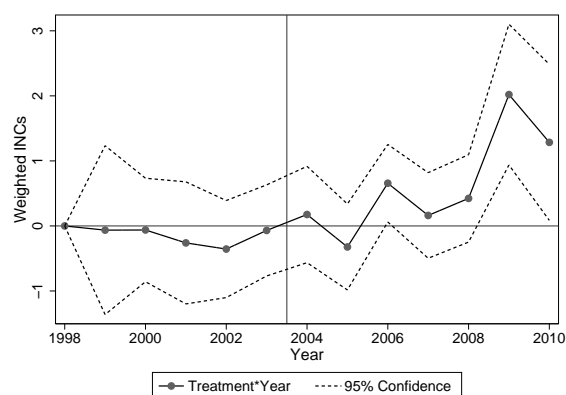
(a) Warnings



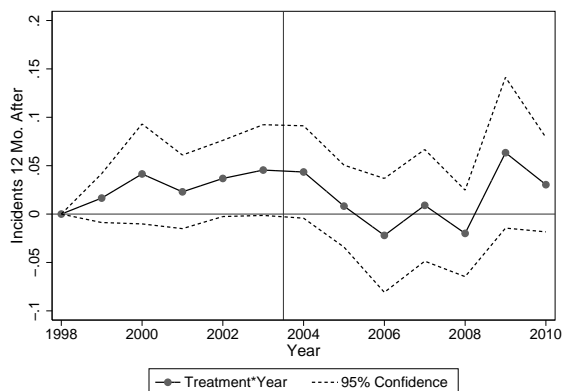
(b) Component Shut-Ins



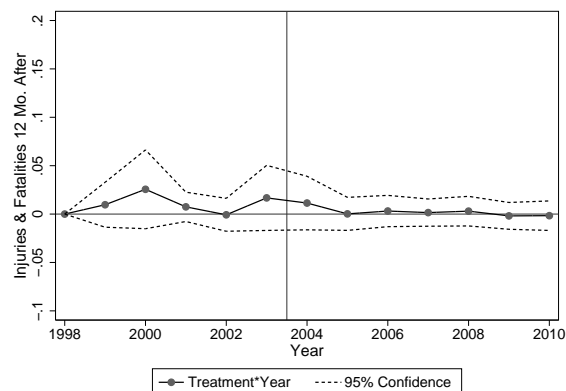
(c) Facility Shut-Ins



(d) Weighted INCs



(e) Incidents 12 Mo. After



(f) Injuries & Fatalities 12 Mo. After

FIGURE 4.— Difference-in-Differences Coefficients for (Treatment * Year Indicators)

5. CONCLUDING REMARKS

There is an extensive literature on the effectiveness of increasing the frequency of inspections and/or inspection resources to improve environmental outcomes. At the same time,

anecdotal evidence and concerns are often raised about the relationship between an inspector and the inspected party and the potential for regulatory capture. Yet, little attention has been given to the allocation and organization of inspection resources. In this paper we examine two parameters important in the design of inspections—the effect of increasing the number of inspectors and the effect of decreasing the familiarity between the inspectors and the inspected party. While we focus on offshore oil platforms, the concern that inspectors are not being optimally employed is ubiquitous across regulatory agencies.²⁸

Given concerns of endogeneity we construct an instrumental variable based on exogenous weather-related changes in the number of inspectors sent to inspect a platform. We also exploit the agglomeration of two inspection offices to obtain an exogenous determinant of the number of times an inspector has visited the same platform. We find more inspectors result in more severe sanctions being issued by the inspectors. While we do not have data on the costs and benefits of inspectors and compliance, our results suggest that at the very least, the level of staffing of inspectors in the Gulf of Mexico was inadequate to detect all violations. We also find that reducing the repeat visits of the same inspector to the same platform results in more severe sanctions being issued. Both of these results are consistent with inspectors undergoing industry pressure not to issue violations.

While we find more violations would be detected if more inspectors were sent on an inspection or familiarity between the inspector and the inspected party is reduced, we do not find evidence that increasing the enforcement actions results in fewer self-reported incidents (such as oil spills, injuries and fatalities). There are several possible reasons for our findings, which are ripe for further research. First, it is possible that the sanctions issued to oil and gas platforms were not stringent enough during the time period of our study to reduce the risk of more accidents. Indeed this was a conclusion made after the BP Deepwater Horizon by the director of the offshore regulatory agency, who stated that “sanctions that are cur-

²⁸For example, see “Internal Audits Reveal Federal Enforcement Lapses,” *Energy and Environment Daily*, April 18, 2011 on the Mine Safety and Health Administration (MSHA) and Tom Zeller, Jr., “Nuclear Agency Is Criticized as Too Close to Its Industry,” *New York Times*, May 7, 2011.

rently available to deter and punish violations of safety and environmental standards and regulations, must be substantially strengthened” [Bromwich, 2010]. Since then, the average inspection in the Gulf has received more INCs, however civil penalties have not increased in frequency or magnitude and attempts to raise the penalties have not been authorized by Congress.²⁹ Second, it is possible that while the number of self-reported incidents is unchanged, the number of actual incidents has decreased. Government enforcement agencies often rely upon self-reports of incidents they would not otherwise detect. Thus, the more stringent enforcement brought about by changing the number or mix of inspectors might induce both deterrence and increased self-reporting. This is plausible since failure to report an incident is itself a law violation, and enhanced scrutiny by inspectors might cause firms to both take better care to comply with the law and to self-report incidents if they happen.

²⁹According to a report analyzing enforcement actions before and after the BP oil spill, prepared by staff members of the House Natural Resources Committee “Dangerous Drillers: Offshore Safety Lapses Continue Three Years After BP Spill,” http://democrats.naturalresources.house.gov/sites/democrats.naturalresources.house.gov/files/2013-05-10_BP_Spill_DangerousDrillers.pdf

REFERENCES

- A.A. Alchian and H. Demsetz. Production, information costs, and economic organization. *The American Economic Review*, 62(5):777–795, 1972.
- Joshua D. Angrist and Guido W. Imbens. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, 90(430):431–442, 1995.
- Gary S. Becker. Crime and punishment: An economic approach. *The Journal of Political Economy*, 76(2):169–217, 1968.
- Gary S. Becker and George J. Stigler. Law enforcement, malfeasance, and compensation of enforces. *Journal of Legal Studies*, 3(1):1–18, 1974.
- Roland Bénabou. Groupthink: Collective delusions in organizations and markets. *The Review of Economic Studies*, 80(2):429–462, 2013.
- Helen Bernhard, Ernst Fehr, and Urs Fischbacher. Group affiliation and altruistic norm enforcement. *American Economic Review*, 96(2):217–221, 2006.
- Roger Bowles and Nuno Garoupa. Casual police corruption and the economics of crime. *International Review of Law and Economics*, 17(1):75–87, 1997.
- Robert Boyd, Herbert Gintis, Samuel Bowles, and Peter J Richerson. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6):3531–3535, 2003.
- Michael R. Bromwich. Implementation Plan In Response to the Outer Continental Shelf Safety Oversight Boards September 1, 2010 Report to the Secretary of the Interior. Technical report, 2010.
- Yan Chen and Sherry X. Li. Group identity and social preferences. *American Economic Review*, 99(1):431–457, 2009.
- Mark Cohen. *Monitoring and Enforcement of Environmental Policy*, volume 3. Edward Elgar publishers, 1999.
- Mark A. Cohen. Optimal enforcement strategy to prevent oil spills: An application of a principal-agent model with moral hazard. *Journal of Law and Economics*, 30(1):23–51, 1987.
- H. Corman and H.N. Mocan. A time-series analysis of crime, deterrence, and drug abuse in New York City. *American Economic Review*, 90(3):584–604, 2000.
- Rafael Di Tella and Ernesto Schargrodsky. Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, (1):115–133, 2004.
- F. Drago, R. Galbiati, and P. Vertova. The deterrent effects of prison: Evidence from a natural experiment. *Journal of Political Economy*, 117(2):257–280, 2009.

- Dietrich Earnhart. Panel data analysis of regulatory factors shaping environmental performance. *Review of Economics and Statistics*, 86(1):391–401, 2004.
- Dennis Epple and Michael Visscher. Environmental Pollution: Modeling Occurrence, Detection, and Deterrence. *Journal of Law and Economics*, 27(1):29–60, 1984.
- Luis Garicano, Ignacio Palacios-Huerta, and Canice Prendergast. Favoritism under social pressure. *Review of Economics and Statistics*, 87(2):208–216, 2005.
- Kishore Gawande and Timothy Wheeler. Measures of Effectiveness for Governmental Organizations. *Management Science*, 45(1):42–58, 1999.
- Monsterrat V. Grau and Theodore Groves. The Oil Spill Process: The Effect of Coast Guard Monitoring on Oil Spills. *Environmental and Resource Economics*, 10(4):315–339, 1997.
- Wayne B Gray and Mary E Deily. Compliance and enforcement: Air pollution regulation in the us steel industry. *Journal of environmental economics and management*, 31(1):96–111, 1996.
- Wayne B. Gray and Jay P. Shimshack. The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*, 5(1):3, 2011.
- B.H. Hamilton, J.A. Nickerson, and H. Owan. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of political Economy*, 111(3):465–497, 2003.
- Catherine Hausman. Corporate incentives and nuclear safety. *American Economic Journal: Economic Policy*, Forthcoming.
- Douglas D Heckathorn. Collective action and the second-order free-rider problem. *Rationality and Society*, 1(1):78–100, 1989.
- James J. Heckman, Sergio Urzua, and Edward Vytlacil. Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432, 2006.
- Eric Helland. The enforcement of pollution control laws: Inspections, violations, and self-reporting. *Review of Economics and Statistics*, 80(1):141–153, 1998.
- Joseph Henrich and Robert Boyd. Why people punish defectors. *Journal of theoretical biology*, 208(1):79–89, 2001.
- Guido W. Imbens and Joshua D. Angrist. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475, 1994.
- International Association of Oil & Gas Producers. Aviation weather guidelines, October 2005.
- Ginger Z. Jin and Jungmin Lee. The imperfection of human inspectors: Lessons from Florida restaurant inspections. 2011.
- E. Kandel and E.P. Lazear. Peer pressure and partnerships. *Journal of Political Economy*, pages 801–817,

1992.

Robert E. Klitgaard. *Controlling corruption*. University of California Press (Berkeley), 1988.

Benoit Laplante and Paul Rilstone. Environmental inspections and emissions of the pulp and paper industry: The case of Quebec. *Journal of Environmental Economics and Management*, 31(1):19–36, 1996.

Steven D. Levitt. Juvenile crime and punishment. *The Journal of Political Economy*, 106(6):1156–1185, 1998.

Steven D. Levitt. Testing the economic model of crime: The National Hockey League’s two-referee experiment. *The BE Journal of Economic Analysis & Policy*, 1(1):2, 2002.

Lance Lochner and Enrico Moretti. Estimating and testing non-linear models using instrumental variables. *NBER Working Paper*, 2011.

Wesley A. Magat and W. Kip Viscusi. Effectiveness of the EPA’s regulatory enforcement: The case of industrial effluent standards. *Journal of Law and Economics*, 33(2):331–60, 1990.

Toni Makkai and John Braithwaite. In and out of the revolving door: Making sense of regulatory capture. *Journal of Public Policy*, 12:61–78, 1992.

Robert E. McCormick and Robert D. Tollison. Crime on the court. *The Journal of Political Economy*, 92(2):223–235, 1984.

Dilip Mookherjee and I.P.L. Png. Corruptible law enforcers: How should they be compensated? *The Economic Journal*, 105(428):145–159, 1995.

Office of Inspector General, Department of the Interior. Outer continental shelf safety oversight board, report to the secretary of the interior ken salazar, 2010a.

Office of Inspector General, Department of the Interior. Investigative report: Island operating company et al, 2010b.

Office of Inspector General, US Department of the Interior. A new horizon: Looking to the future of the bureau of ocean energy management, regulation and enforcement, December 2010.

Offshore Energy and Minerals Management. Gulf of Mexico OCS region field operations efacilities inspections team report, March 2009.

Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge university press, 1990.

A Mitchell Polinsky and Steven Shavell. chapter Public enforcement of law. Edward Elgar Publishing, 2009.

Nancy L Rose. Profitability and product quality: Economic determinants of airline safety performance. *Journal of Political Economy*, pages 944–964, 1990.

Susan Rose-Ackerman. *Corruption and Government: Causes, Consequences, and Reform*. Cambridge University Press, 1999.

- Jay P. Shimshack and Michael B. Ward. Regulator Reputation, Enforcement, and Environmental Compliance. *Journal of Environmental Economics and Management*, 50(3):519–540, 2005.
- Doug Slitor. Measuring safety and compliance in the U.S. offshore oil and gas industry. In *SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production*, 2000.
- James Stock and Douglas Staiger. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586, 1997.
- U.S. Army Coastal Engineering Research Center. *Shore Protection Manual*. Waterways Experiment Station, Fort Belvoir, VA, 4th edition, 1984.
- W Kip Viscusi. The impact of occupational safety and health regulation, 1973-1983. *The RAND Journal of Economics*, pages 567–580, 1986.
- Toshio Yamagishi. The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51(1):110, 1986.

APPENDIX A: EXAMPLE FROM THE POSSIBLE INCIDENT OF NON-COMPLIANCE (PINC) LIST

In the PINC list, there are guidelines governing many aspects of oil and gas operations including testing and checking specific components' functionality, personnel training, record-keeping, drilling operations, and oil spill response plans. Below is an example of a general environmental protection guideline listed in the PINC list.³⁰

Does the operator ensure activities authorized in this part are carried out in a manner that provides for protection of the environment?

Inspection procedure:

Visually check the waters surrounding the facility when en route to, approaching, departing, or passing a facility by boat, helicopter, or fixed wing aircraft. Look for sheens, slicks, wastes, and other pollutants originating from the facility, risers, or pipelines. During the facility inspection, check the water for signs of any pollutants.

If noncompliance exists

- *Issue a warning (W) INC if:*

The cause of the discharge has been identified and corrected prior to the inspection or observation, and no further pollution is occurring.

- *Issue a component shut-in (C) INC if:*

A specific component has been determined to be the cause of pollution and the pollution is ongoing at the time of the inspection or observation.

- *Issue a facility shut-in (S) INC if:*

More than one specific component has been determined to be the cause of the pollution and the pollution is ongoing at the time of the inspection or observation.

The PINCs outline the enforcement actions to be taken upon different degrees of non-compliance. The inspector *must* take these actions; however, one might imagine that there could be circumstances when the interpretation of the definition of a specific component or the timing of the incident might influence the enforcement action taken.

³⁰For the full list of PINCs, see:
http://bsee.gov/uploadedFiles/BSEE/Enforcement/Inspection_Programs/OFFICEPINCLIST_JBA%20Changes%20Included.pdf

APPENDIX B: WHY ARE THE IV ESTIMATES LARGER THAN THE OLS ESTIMATES?

As Table IV in the paper illustrates, the IV estimates on the impact of inspector group size on enforcement actions are always larger than the corresponding OLS estimates. Besides endogeneity, there could be two other reasons driving this result: measurement error in group size and non-linearity in the effect of group size. As discussed in section 4.1, measurement error in the number of inspectors is not as likely as non-linearity in the effects, and therefore we examine this in more detail. The IV estimates use variation in the inspection group size induced by wind speed, whereas OLS estimates use all the variation. If weather affects a particular level shift in the group size and this level shift has a larger effect on enforcement, then the IV and the OLS estimates will differ. We find that the instrument has the largest effect in increasing the number of inspectors from one to two (Table X, column 1), less of an effect in moving the number of inspectors from two to three (column 2), and no statistically significant effect in moving from three to four or more (column 3).

To examine whether there are non-linearities in the enforcement actions by inspection group size, we estimate the OLS specifications in Table IV using dummies for each level of group size instead of the count of inspectors. This approach gives us a separate coefficient on each group size level: moving from one to two or more, three or more, and four or more inspectors (Table XII). In most cases, F tests reject the null that the coefficients are equal. Therefore, it appears that the true relationship between the enforcement actions and the number of inspectors is non-linear, but because we do not have enough instruments, we have estimated a linear model and restricted the effect to be constant. Weights for IV estimators in the presence of non-linearity have been derived by Angrist and Imbens [1995], Heckman et al. [2006] and Lochner and Moretti [2011]. Following Lochner and Moretti [2011] we re-weight the OLS estimates of in Table XII using estimated IV weights, and use a general Wald test, robust to a non-linear relationship between the endogenous regressors and the dependent variable, to test for exogeneity.³¹ The re-weighted OLS coefficients are larger than the OLS estimates (Table XII), however the re-weighted OLS coefficients are still smaller than the IV estimates. Using the general Wald test, we cannot reject the null hypothesis of exogeneity. Therefore, we have some confidence in the OLS estimates, however, the IV estimate, which is the weighted average of the marginal effects is still useful for guiding inspection policy.

³¹If the true relationship between the number of inspectors and enforcement is non-linear, then the traditional Hausman test is uninformative.

TABLE X
INSTRUMENT'S EFFECT ON LEVEL SHIFTS IN GROUP SIZE

	(1) More than 1	(2) More than 2	(3) More than 3
Std. Dev. WSPD	.041** (.020)	.021** (.008)	.001 (.001)
Controls	Yes	Yes	Yes
n	2,008	2,008	2,008
Mean of Dep. Var.	.30478	.05129	.00697

Notes: Each column represents a separate OLS regression of variation in wind speed on indicators for whether there were more than 1, 2, or 3 inspectors in the group. All regressions include the same controls as in previous regressions. Sample includes production inspections in Lake Jackson between October 2003 and September 2010. Robust standard errors are clustered by operator (for 91 operators). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

TABLE XI
ESTIMATES AT EACH LEVEL INCREASE IN GROUP SIZE

	(1) Warnings	(2) Component	(3) Facility	(4) Weighted
Two Inspectors	.011 (.061)	.098 (.062)	.017 (.016)	.273 (.172)
Three Inspectors	.259 (.176)	.645*** (.237)	.119** (.050)	2.025*** (.600)
Four or More Inspectors	.057 (.352)	.582 (.386)	.030 (.090)	1.338 (1.288)
Controls	Yes	Yes	Yes	Yes
n	2,008	2,008	2,008	2,008
Mean of Dep. Var.	.45518	.48257	.03386	1.55578

Notes: Each column represents a separate OLS regression of indicators for the size of the group on the enforcement actions (W, C, F, and W+2C+4F). All regressions include the same controls as in previous regressions. Sample includes production inspections in Lake Jackson between October 2003 and September 2010. Robust standard errors are clustered by operator (for 91 operators). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

TABLE XII
RE-WEIGHTED OLS COEFFICIENTS

	Warnings		Component		Facility		Weighted	
	(1) OLS	(2) ROLS	(3) OLS	(4) ROLS	(5) OLS	(6) ROLS	(7) OLS	(8) ROLS
No. Inspectors	.069 (.054)	.101 (.057)	.216*** (.053)	.263*** (.084)	.036** (.015)	.049** (.019)	.644*** (.153)	.823*** (.241)
General Wald (p-value)	.910		.696		.050		.402	

Notes: OLS estimates are the same as found in Table IV. Re-weighted OLS estimates from derived by using IV weights to calculate an average of the coefficients in Table . The General Wald test [Lochner and Moretti, 2011] compares the re-weighted OLS in this table to the linear IV estimator (found Table IV).