

The Role of Organizational Scope and Governance in Strengthening Private Monitoring

**Lamar Pierce
Michael W. Toffel**

Working Paper

11-004

August 22, 2012

Copyright © 2010, 2011, 2012 by Lamar Pierce and Michael W. Toffel

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

The Role of Organizational Scope and Governance in Strengthening Private Monitoring*

Lamar Pierce
Olin Business School
Washington University in St. Louis
St. Louis, MO 63130
pierce@wustl.edu

Michael W. Toffel
Harvard Business School
Morgan Hall 497
Boston, MA 02163
mtoffel@hbs.edu

August 22, 2012

Governments and other organizations often outsource activities to achieve cost savings from market competition. Yet such benefits are often accompanied by poor quality resulting from moral hazard, which can be particularly onerous when outsourcing the monitoring and enforcement of government regulation. In this paper, we argue that the considerable moral hazard associated with private regulatory monitoring can be mitigated by understanding conflicts of interest in the monitoring organizations' product/service portfolios and by the effects of their private governance mechanisms. These organizational characteristics affect the stringency of monitoring through reputation, customer loyalty, differential impacts of government sanctions, and the standardization and internal monitoring of operations. We test our theory in the context of vehicle emissions testing in a state in which the government has outsourced these inspections to the private sector. Analyzing millions of emissions tests, we find empirical support for our hypotheses that particular product portfolios and forms of governance can mitigate moral hazard. Our results have broad implications for regulation, financial auditing, and private credit- and quality-rating agencies in financial markets.

* We thank Melissa Ouellet for excellent research assistance and Ryan Buell, Robert Huckman, Donald Palmer, Daniel Snow, Jason Snyder, and Andrew von Nordenflycht for helpful comments and suggestions. We gratefully acknowledge financial support from the Division of Research at Harvard Business School, the Harrington Foundation at the University of Texas at Austin, and the Olin Business School at Washington University in St. Louis.

The Role of Organizational Scope and Governance in Strengthening Private Monitoring

Governments have long debated which societal functions should be outsourced to private firms. Often motivated by potential cost reductions from market competition (Williamson 1985), outsourcing services to the private sector also risks moral hazard, which can reduce service quality (Sclar 2000; Levin and Tadelis 2010; Becker and Milbourn 2011). Considerable theory and empirical analysis have shown that, without costly oversight, outsourcing functions to firms can lead to agency problems due to the different incentives of principals and agents (Jensen and Meckling 1976; Klein, Crawford, and Alchian 1978; Reichelstein 1992). Yet governments are increasingly outsourcing services they have traditionally performed (Freeman and Minow 2009), including garbage and recycling collection, fire departments, prisons, and utilities such as electricity, water, cable, and Internet (Hart, Schleifer, and Vishny 1997; Cabral, Furquim, and Lazzarini 2010; Seamans 2011). Even military functions, which Williamson (1999) argued must necessarily be provided by government, have been outsourced both historically (e.g., Nepalese Gurkhas and private Italian armies) and increasingly in modern times (e.g., Blackwater and Afghan warlords) (Minow 2005; Baum and McGahan 2009).

Some governments have also outsourced the monitoring of compliance with laws and regulations. Examples include private arbiters (Richman 2004), auditors (Corona and Randhawa 2010), certified public accountants (Moore et al. 2006), environmental monitors (Seifter 2009), credit-rating agencies (He et al. 2011), stock exchanges (Jamal 2008), and retailers enforcing age limits for alcohol and tobacco sales. This outsourcing of monitoring is similar to formal three-tiered agency models in accounting and economics, where the principal (government) hires a supervisor (private monitor) to monitor the behavior of the agent (regulated entity) (Antle 1984; Tirole 1986). In these models, much of the efficiency gain from hiring a supervisor to monitor the agent is compromised by the propensity of the agent to buy the supervisor's collusion through side payments. In financial auditing, for example, such conflicts of interest are known to generate fraud (e.g., Khalil and Lawaree 2006) and are exacerbated when regulated entities

are allowed to choose their own monitors (Boyd 2004). Such customer choice may be socially beneficial by helping monitors build specialized client-based knowledge and by allowing customers to more efficiently choose convenient and lower-priced monitors. But private monitors' desire to please customers in order to solicit future business results in an "arrangement [that] threatens to punish [private monitors] who might otherwise be inclined to do an A+ rather than a D+ job," resulting in substantially more leniency than government monitors might show (Seifter 2009: 99, 103). This moral hazard problem is similar to that of corruption in government officials who ignore legal or regulatory violations for bribes or political favors (Bertrand et al. 2007; Dal Bó 2006; Shleifer and Vishny 1993; Laffont and Tirole 1993; Fan, Lin, and Treisman 2009).

In this paper we argue that one solution to reducing the moral hazard problem in private monitoring lies in understanding how a monitor's incentives to provide leniency are influenced by the scope of its business activities and by its governance structure. While prior work has focused on the incentive of the monitoring activity's direct profitability (e.g., Becker and Milbourn 2011; Bolton et al. 2012), we focus on the monitoring firm's incentive of cross-selling to the monitored party. Although some firms specialize exclusively in private monitoring, many firms—including financial auditors, vehicle emissions inspectors, and law firms—operate in additional markets and must consider the impact of their monitoring stringency on the profitability of their other products and services. The perverse incentives to provide leniency in order to cross-sell were recently highlighted during the Arthur Andersen/Enron scandal. Prior to the reforms of the Sarbanes-Oxley Act (2002), firms could charge below-cost prices for auditing to bolster their cross-selling of more lucrative consulting services (Levitt 2000). For the audited firm, paying for these consulting services could serve as a side payment for auditing leniency.

We argue that the private monitoring market functions like three-tiered principal-supervisor-agent models developed by Tirole (1986) that have been used to explain leniency in financial auditing (Khalil and Lawaree 2006). In private monitoring, the agent (monitored party) pays the supervisor (monitor) for providing oversight on behalf of the principal (government). The agency problem in this setup is that the

agent might give the supervisor a side payment to encourage leniency. Given that profitable cross-selling contracts can act as side payments for leniency, just as lucrative management consulting contracts were thought to serve as side payments for lenient financial auditing (Levitt 2000), we argue that a monitoring firm's portfolio of cross-sellable products and services is a primary predictor of leniency in the monitoring activity.

Greater leniency is especially likely when it can be traded for large side payments, as is the case in markets characterized by long-term customer loyalty with repeated high-margin cross-sales. But even in such cases, the monitor must also consider the extent to which providing leniency will create reputational spillovers of dishonesty that can erode its sales (Nickerson and Silverman 2003; Mayer, Nickerson, and Owan 2004). If customers of the cross-sold product are vulnerable to moral hazard that can lead to unanticipated poor product quality, they may fear that the same dishonest firm that is leniently monitoring them will also dishonestly cross-sell them low-quality products. We argue that the extent of moral hazard risk is determined by the inherent quality uncertainty of the cross-sold product and by the frequency of the transactions that form a long-term customer relationship. The efficacy of monitoring can therefore be improved by outsourcing monitoring to firms whose scope does not include activities with strong profit opportunities and low moral hazard risks associated with cross-selling.

We argue that firm governance structure also plays a critical role in predicting monitoring leniency, affecting not only individual managerial incentives but also corporate incentives to manage the risk of reputation loss and regulatory sanctions. Consistent with the literature on franchising and managerial control (LaFontaine and Shaw 2005), independent owners of regulatory monitors have strong incentives for leniency because they profit directly from it and have no corporate parent to monitor their behavior. Managers at wholly owned subsidiaries, however, have weaker incentives to improve local profits through leniency (Bradach 1997); they cannot claim facility profits yet still risk criminal charges or the loss of license. Furthermore, leniency may also hurt a corporation's brand equity and reputation with the government, which provides an incentive for the corporation to increase its oversight of local operations to assure stringent monitoring of customers (Williamson 1983). We also propose that branded

franchising is another governance form that can affect leniency. While franchisees may have strong incentives for leniency, the franchisor has strong competing incentives to police franchisees in order to avoid harm to the brand's reputation with the government.

We explore these issues in the context of automobile emissions testing. Many state governments license firms to monitor the regulatory compliance of vehicles. Although these firms must appear legitimate to the agencies that license them, they have strong incentives to relax their monitoring because passing vehicles significantly increases the likelihood that customers will return (Hubbard 2002). We examine a panel of 2.7 million vehicle inspections conducted in the five-year period of 2000 through 2004 by 3,500 private-sector inspection facilities in the New York metropolitan area. We first confirm that car owners in our sample, like those already studied in California (Hubbard 2002), are less likely to return to facilities that fail their vehicles. We then test our hypotheses and find considerable differences in leniency across firms of differing activity scopes and governance structures. In terms of firm scope, service and repair facilities and dealerships—firms that cross-sell high-margin products and services to loyal customers—exhibit more leniency than do gasoline retailers, whose cross-sold product is low-margin gasoline and whose customers feel less loyalty. With respect to governance, branded and subsidiary facilities, which have internal governance structures designed to monitor activities, are consistently less lenient than independent facilities. Our results also suggest that cross-selling profitability and the potential for reputational spillovers predict firm's strategic decisions on monitoring leniency. Our analysis is consistent with a recent investigation in New York State (New York State Department of Environmental Conservation 2010) that cited 40 facilities for fraudulent inspections, all but two of which were independent unbranded repair facilities.

Our paper addresses the growing need for research that integrates private and public interests (Mahoney, McGahan, and Pitelis 2009). A rapidly expanding empirical literature is examining the competitive and regulatory impact of the blurring of traditional boundaries between public and private interests (Cabral et al. 2010, 2012; Seamans 2011), but little is known about how a firm's characteristics will affect its conduct of traditionally public activities. Furthermore, the large literature on financial

auditing has few parallel studies from other industries that also employ private firms to provide public monitoring activities (e.g., Lennox and Pittman 2010). Our paper provides a rare complement to this literature in an industry that, like auditing, has economic and social importance, but has considerably greater variety in the market scope of monitors. Furthermore, we build on several previous studies of vehicle emissions testing that establish the existence of and incentives for fraudulent leniency (Hubbard 1998, 2002; Oliva 2012; Pierce and Snyder 2008).

Our results also have implications for policymakers and managers. Our results indicate that governments, when licensing facilities and when targeting their investigations of licensed facilities, should carefully examine how private monitors' other lines of business can create perverse or beneficial incentives for monitoring stringency. Our results also suggest that, compared to the potential efficiency gains that often attract governments to market-based solutions, the actual gains associated with privatizing monitoring may be limited by the widespread incentives for such monitors to provide lenient oversight. For managers, our results suggest license assignment to particular types of firm; namely subsidiaries, branded franchisees, and gasoline retailers. Such firms can make an argument for the reduced likelihood of malfeasance under their monitoring

Theory and Hypotheses

In a market for private regulatory monitoring, for-profit firms can face conflicting incentives regarding the stringency of monitoring. They often operate under a government license that requires stringent monitoring, with the consequences for noncompliance ranging from financial penalties to loss of license to civil and criminal penalties. But contrary incentives may arise from customers seeking lenient private monitors to avoid the costs resulting from the detection of infractions, a process referred to in the accounting literature as “audit shopping” (e.g., Davidson et al. 2006). This demand for leniency creates a situation similar to the situation in Tirole's (1986, 1992) three-tiered principal-supervisor-agent models, where within a firm, an intermediary supervisor (the monitor) engages in side contracts with an employee (the agent) instead of serving the senior manager or owner (the principal). These models have been

applied to the financial auditing industry (Khalil and Lawaree 2006) and to bureaucratic corruption (Laffont and N'Guessan 1999).

In private regulatory monitoring, which also spans three parties, the firm (the monitor), licensed by the state (principal), may ignore or downplay observed violations and profit from side payments from the party it is charged with monitoring (the agent). Side payments in exchange for leniency are believed to be more prevalent when there is competition amongst monitoring firms, as is the case with bond ratings (Becker and Milbourn 2011) and corruption (Laffont and N'Guessan 1999). Recent theoretical models by Drugov (2010) show that, when there is competition, side payments (bribes, in his models) are smaller on average but more common than when there is no competition. In many privatized monitoring markets in which the regulator sets a standard price—including the market for vehicle emissions testing in many states—leniency can become a critical basis for competition, along with location, scheduling availability, and service quality.

Understanding which types of firms are particularly prone to leniency can help governments target their necessarily limited oversight. Below, we argue that organizational scope can affect a firm's incentive for leniency because the cross-sale of other profitable products and services can serve as side payments for lenient monitoring. We argue that leniency is especially likely under two conditions that generate large side payments for monitors: first, when cross-selling opportunities are profitable and frequent, and second, when customers face little moral hazard risk from the combination of (a) uncertain quality in the cross-sold product and (b) short-term relationships. We also argue that two forms of private governance—corporate ownership and brand affiliation—create incentives for corporations to police local operations to assure stringent monitoring. In sum, we posit that the organizational scope and private governance of private monitors will predict the likelihood of illicit leniency.

Organizational Scope and Incentives for Leniency

Many markets feature a mix of private monitors, with some entirely specialized in the monitoring activity and others operating as multi-market firms. The latter have an opportunity to trade monitoring

leniency for the implicit side payment of buying the firm's other goods and services. The expected profitability of such trades is driven by two factors associated with cross-selling: the monitor's profit opportunity and the monitored party's risk of moral hazard. The profit opportunity determines the value from potential cross-sales and the risk of moral hazard determines the likelihood of future cross-sales.

The profit opportunity from extending leniency depends on the profitability and frequency of cross-sold transactions. The opportunity to cross-sell a large, high-margin product or service creates incentives for firms to exchange leniency for an implicit side-payment of such purchases. While the recipients of leniency cannot be formally obligated to buy other products in the future, this form of implicit side payment is legally safer than an explicit bribe. The value of this side payment is also increased by the frequency of future sales of the cross-sold product. If cross-selling to the monitored party offers a long-term stream of profitable transactions, the incentive to capture that through leniency is even greater. Under these conditions, the monitor must consider the impact of its monitoring activity on the potential for a long-term sales relationship with the monitored party. For example, when integrated firms that conduct both financial auditing and management consulting are auditing to enforce accounting standards, concern for overall profitability encourages them to consider how their monitoring stringency affects their opportunities to cross-sell large, high-margin, and repeated consulting services (Levitt 2000). Similarly, investment banks that issue equity recommendations are likely to consider the impact of this analysis on future fees from merger-and-acquisition deals.

Hypothesis 1a: Private regulatory monitoring establishments will be more lenient when they face profitable opportunities to cross-sell to repeat customers.

Even when cross-selling provides a profitable opportunity for the monitor, the monitored party might be wary of the cross-sold product due to moral hazard risk. This risk arises from information asymmetry regarding the quality of the cross-sold product, which creates the possibility that the firm might misrepresent product quality to the customer. From the customer's point of view, then, a private monitor dishonest enough to trade leniency for the implicit side payment of cross-selling might also be dishonest enough to exaggerate the quality of its other products or services. For example, a patient might

be happy to have a doctor falsely sign an immunization form for his or her child, but might think twice if advised by the same doctor to undergo an expensive and risky procedure. Fear of such quality deception can substantially increase the cost of the side payment to the monitored party, which is the price they pay for the cross-sold product minus the value they receive for it. In particular, fear of quality deception erodes the value of the cross-sold product (without affecting its price). As with standard moral hazard problems, this would make monitored parties less likely to buy the cross-sold product, thereby reducing expected profitability for the private monitor seeking to sell it. In short, reputational spillover from monitoring to the cross-sold product could create an umbrella brand of dishonesty across a firm's products and services (Wernerfelt 1988), similar to reputational spillovers found in past studies (Jensen 1992; Nickerson and Silverman 2003; Mayer, Nickerson, and Owan 2004; Mayer 2006; Bénabou and Tirole 2006).

The monitored party faces greater moral hazard risk from the cross-sold products and services when quality is unobservable ex ante and when transactions are infrequent. For cross-sold products and services whose quality is observable, monitoring leniency should pose little threat to the cross-selling opportunity. But with experience goods, for which quality is observable only after use (Nelson 1970), or with credence goods, for which quality is unobservable even after use (Darby and Karni 1973), customers may rightfully fear opportunistic behavior (Emons 1997). Such moral hazard concerns are attenuated, however, when firms expect long-term relationships consisting of repeated transactions that might be endangered by moral hazard (Holmstrom 1979; Williamson 1985). Therefore, when moral hazard exists in the cross-sold market due to unobservable quality and infrequent transactions, monitored parties will be less likely to trade purchases of these products for monitoring leniency. Any firm willing to dishonestly help them for profit will be expected to also dishonestly hurt them for profit.

Hypothesis 1b: Private regulatory monitoring establishments will be less lenient when they have cross-selling opportunities with moral hazard risk that is due to uncertain quality and low transaction frequency.

We present the nexus of Hypotheses 1a and 1b in Figure 1, a two-dimensional plot that shows our prediction of an increase in leniency when (a) customers perceive less risk of moral hazard and (b) monitoring firms faces larger profit opportunities from cross-selling. Higher levels of predicted leniency are represented with darker shading.

INSERT FIGURE 1 HERE

Private Governance and Corporate Oversight

Monitoring firms risk being expelled from the market if government investigations detect lenient monitoring. This risk constrains leniency for all monitoring firms to some extent, but the potential cost of government sanctions differs across firm along two dimensions: ownership and branding. We summarize our argument here and provide more details below. The monitoring company's ownership affects the consequences of potential government sanctions because those monitoring establishments that are subsidiaries create legal liability not only for themselves (as would be the case for independently owned establishments) but also for their parent companies. Ownership also affects leniency because the incentives to increase profits through leniency are weaker for subsidiary managers than they would be for independent owners. The second dimension, branding, affects the consequences of potential government sanctions because the actions of one branded establishment can draw regulatory attention to other establishments that share its brand. These dimensions are not mutually exclusive and can simultaneously affect monitoring leniency. A market can include unbranded subsidiaries, branded subsidiaries (corporate-owned chains), branded independents (franchises), and unbranded independents.

A monitor's ownership affects its leniency because while an independent, single-location firm risks the loss of just its own monitoring license, a subsidiary of a multi-unit or multi-location firm risks government sanctions that impose additional costs on the parent company and its other establishments. Locations and business units that did not directly benefit from leniency might nevertheless receive increased government scrutiny, which is not only costly in its own right, but also risks revealing other violations. This increased risk is not merely a result of the firm's scale of monitoring activities; other

corporate business units engaged in other activities might nevertheless be impacted by increased government monitoring. A broad literature shows that facility compliance efforts are indeed influenced by inspections and enforcement activities targeted at other firms (Epple and Visscher 1984; Cohen 1987, 2000; Shimshack and Ward 2005; Short and Toffel 2008; Thornton et al. 2005).

One might imagine that subsidiaries, compared to independently owned firms, have greater access to legal resources with which to defend themselves against government fraud charges and that this could reduce their risk of being investigated. However, this factor seems unlikely to deter government investigations and prosecution because elected officials and prosecutors often derive political benefits from targeting larger entities. Thus, compared to independently owned firms, subsidiaries likely perceive a heightened risk of being investigated, which further increases the incentives for corporate managers to police their subsidiaries in order to deter leniency.

In addition, managers of subsidiaries have weaker incentives to engage in leniency than managers of independent firms do. Many independently owned firms are managed by their owners, who face strong incentives to maximize profits. Even when independently owned firms are managed by someone else, the owner is typically local and owns only a few locations, which minimizes the cost to the owner of ensuring that the manager's behavior is consistent with the owner's high-powered incentives. In contrast, subsidiary managers are typically given low-powered incentives by the owners (Bradach 1997). These low-powered incentives are often due to agency concerns that managers will sacrifice long-term investments for short-term profits (Wulf 2002) or will overallocate effort toward the tasks on which incentives are based (Holmstrom and Milgrom 1991). Low-powered incentives might also be efficient for subsidiary managers due to hazards from asset specificity or the need to adaptively coordinate (Williamson 1985). Even if subsidiary managers are compensated in part based on the establishment's performance, their incentives will be inherently weaker than those of independent owners who are the residual claimants of all profits. Subsidiary managers would therefore reap less of the benefits of leniency, while being just as vulnerable to the consequences from government detection, such as the risk of being fired. LaFontaine and Shaw (2005) found that firms with a greater need to protect their brand are

more likely to avoid the high-powered local incentives of franchises and instead control local behavior through corporate ownership.¹ With increased oversight by the corporate parent and less financial incentive to improve performance, managers of subsidiaries are less likely to engage in leniency.

Hypothesis 2a: Private regulatory monitoring establishments that are subsidiaries will be less lenient than those that are independent.

Some monitoring establishments that are not wholly owned subsidiaries are associated with corporations through branding and franchise relationships. Although independent monitoring firms might earn reputations for leniency that attract customers, it is difficult for a brand to do so because any reputation that transcends one location is likely to attract attention from regulators. Furthermore, as noted earlier, positive reputational spillovers from one branded location to others require customers to openly discuss their own solicitation of leniency across their geographically distant social network. Modern Web-based review systems are highly unlikely to transmit explicitly illicit information for fear of detection and enforcement by authorities. Recent work by Jin and Leslie (2009) suggests that reputation with customers across chains and branded franchises may motivate quality improvement among restaurants, but even in that market, it is the government's safety and quality grades that primarily influence firm behavior. Brand owners, therefore, see little upside to an image of leniency, while the downside is very real. A branded establishment caught providing lenient monitoring might invite brand-wide investigations by the state regulator, which might believe that brand-level operating processes are part of the problem, either through weak process control or because the brand owner intentionally selects franchisees prone to leniency.

Branded establishments that sully their brand reputation with regulators risk triggering investigations of their brands' other establishments, just as subsidiaries that sully their owner's reputation risk triggering government investigations of their corporate family. Therefore, branded companies have incentives to carefully select franchisees averse to leniency and to oversee their monitoring activities.

¹ Local managers of subsidiaries and owners of independent firms are likely to have equivalent information on employee behavior, unless managerial turnover is high. In the latter case, leniency may be higher in subsidiaries if employees were accepting direct bribes from customers, although undercover operations in California found this to be relatively rare (Hubbard 1998). Any employee-monitoring advantages in independent facilities are unlikely to counteract these facilities' strong owner incentives for leniency.

Indeed, corporations that franchise their brands often distribute to franchisees an operations manual that serves as “a functional tool for enforcing system standards” (Brams 1999: 77) and often include in their franchise agreements the right to periodically inspect franchisees to verify adherence to these operational standards (Brams 1999; Perkins, Yatchak, and Hadfield 2010). As Lafontaine and Blair (2009: 381) note, a required component of a franchise relationship, according to the U.S. Federal Trade Commission, is that “the franchisor must exert significant control over the operation of the franchisee or provide significant assistance to the franchisee.” Just as the threat of regulatory inspections bolsters firms’ compliance with regulatory requirements (Laplante and Rilstone 1996), the threat of corporate inspection—and of forfeiture of the franchise—should bolster franchisees’ compliance to franchise standards. These dynamics suggest the following hypothesis:

Hypothesis 2b: Private regulatory monitoring establishments affiliated with a multi-location brand will be less lenient than independent monitoring establishments.

It is important to note that we characterize ownership structure and branding as separate but correlated characteristics. While many subsidiaries share a common brand, some do not. Similarly, franchises often share a common brand but not a common owner. Since our theory suggests that ownership structure and brand are both likely to reduce leniency, branded subsidiaries ought to be the least lenient, whereas non-franchised, independently owned firms ought to be the most lenient.

We hypothesize that subsidiary status reduces leniency through two mechanisms: weaker managerial incentives and the risk of negative reputation spillovers in the eyes of the regulatory agency. Brandedness, however, reduces leniency only through efforts by the brand owner to protect the brand’s reputation, since owners of (non-subsidiary) branded facilities—that is, franchisees—have incentives to exhibit leniency given that they are the residual claimant on profits from long-term customer loyalty. Such franchisees are likely to attempt to free-ride on a brand’s reputation (Jin and Leslie 2009) and provide leniency despite the risk to the brand. Similarly, any possible positive reputational spillovers that might motivate increased leniency are likely to occur across establishments that share a brand, rather than across subsidiaries of a common owner, since customers are unlikely to recognize common ownership in

the absence of branding. We therefore expect brandedness to attenuate leniency to a lesser extent than subsidiary ownership structure does.

Hypothesis 3: Affiliation with a multi-location brand will reduce private regulatory monitoring establishments' leniency less than subsidiary structure will.

Empirical Setting

We test our hypotheses in the empirical context of the vehicle emissions testing market, where federal environmental protection regulations require many states to restrict the levels of air pollutants produced by personal and commercial vehicles. Motor vehicle emissions are a major source of air pollution. Transportation accounts for as much as 10% of fine particulate matter emissions in the United States (U.S. Environmental Protection Agency 2007) and, in metropolitan areas, emits nearly half of the total emissions of six heavily regulated “criteria air pollutants,” which include carbon monoxide, particulate matter, ground-level ozone, and nitrogen oxides (Ernst, Corless, and Greene-Roesel 2003). In cities with poor air quality, vehicles account for 35% to 70% of ozone-forming emissions and at least 90% of carbon monoxide emissions (U.S. Environmental Protection Agency 1994). Vehicle emissions inspection and maintenance (I/M) programs can reduce these emissions by 5% to 30% (U.S. Environmental Protection Agency 1994). In our focal state, New York, every registered vehicle built since 1981 and weighing less than 8,500 pounds must be tested annually for emissions of hydrocarbons (HC), carbon monoxide (CO), and nitrogen oxides (NO_x). Vehicles with emissions levels exceeding the legal limits for any of these pollutants—by no matter how little—fail the test. Until 2005, all eligible vehicles received dynamometer tests, which measure pollutants expelled from the vehicle’s exhaust pipe.² All technicians conducting emissions tests must be certified by the New York State Department of Motor Vehicles, which requires (a) at least one year of vehicle repair experience or a diploma from a motor vehicle vocational school and (b) completion of an inspection certification training program, including

² Beginning in 2005, New York introduced an on-board diagnostic system called OBD-II for new models in urban areas, but none of these tests are in our sample. Dynamometer tests continued to be used for older vehicles and for the rest of the state.

passing a written test (New York State Department of Motor Vehicles 2004, 2011). State regulations stipulate equipment specifications, require all testing facilities to purchase standardized equipment from a state-approved vendor, and regulate and enforce standardized equipment maintenance procedures (New York State Department of Environmental Conservation 2004a, 2004c, 2011).

Incentives for Leniency in Private Emissions Monitoring

In the United States, many state governments seeking economic efficiencies have outsourced monitoring of vehicle emissions standards to the private sector, despite potential conflicts of interest between (a) governments, which desire stringency through accurate monitoring, and (b) firms and vehicle owners, who stand to benefit from leniency in the form of fraudulently inaccurate monitoring (National Research Council 2001). Concerns about corruption, collusion, and inaccurate monitoring date back to the 1970s, when state governments began to require periodic vehicle-safety checks and emissions testing and debated whether to establish government-operated facilities or outsource to the private sector (Rule 1978; Lazare 1980). The traditional argument for privatization was one of market efficiency—drivers could conveniently get tested at a local business with strong incentives for efficiency and quality. The argument against privatization was environmental: Repair facilities had so many incentives to build and maintain long-term relationships that they were unlikely to engage in stringent emissions testing (Voas and Shelley 1995; Harrington and McConnell 1999).

Many firms do have strong incentives to relax their monitoring and show leniency to core customers. Hubbard's (2002) analysis of several thousand vehicle inspections in the early 1990s in Fresno, California found that a car owner was significantly more likely to return to an inspection facility that passed his or her vehicle than to one that failed it. As the California Bureau of Automotive Repair (BAR) noted, "it appears, based on BAR enforcement cases, that some stations improperly pass vehicles to garner more consumer loyalty for delivering to consumers what they want: a passing Smog Check result" (California Bureau of Automotive Repair 2011: 22). Any facility unwilling to adjust results to pass a vehicle may lose that customer's immediate and future business—both emissions testing and other

products and services. Owners of noncompliant vehicles have strong incentives to choose lenient facilities and to leverage their patronage to motivate such behavior. Mounting evidence of lenient private monitoring in the vehicle emissions testing market suggests that concerns about inspection fraud and collusion between vehicle owners and inspectors are justified, given that 20% to 50% of noncompliant cars are fraudulently passed, based on estimates from separate samples in California, Mexico City, and New York (Hubbard 1998; Oliva 2012; Pierce and 2012).

Technically, dynamometer-based testing offers ample opportunities for inspectors to fraudulently pass a vehicle, as evidenced by an Atlanta trio who fraudulently passed over 1,400 vehicles over a five-month period in 2011 (Crosby 2011). Not only do vehicles get two chances to pass the test, but inspectors can stop either test if they perceive a problem. This allows them, when a vehicle appears to be failing, to make illegal temporary adjustments such as introducing fuel additives (e.g., denatured alcohol), adjusting the tailpipe probe, or diverting exhaust before it reaches the tailpipe. While these adjustments have become more difficult due to improved testing regulations, other fraudulent techniques remain common. Inspectors can also use a device that simulates a tachometer, thereby allowing the car to test at fewer revolutions per minute (New York State Department of Environmental Conservation 2010). Technicians can also mask emissions problems by shifting the vehicle into the wrong gear during a test, by racing the engine to get the catalytic converter hotter than its normal operating condition, and by entering incorrect vehicle parameters to generate more lenient emissions thresholds (California Bureau of Automotive Repair 2011: 47). An inspector can even substitute a vehicle capable of passing in place of a failing vehicle in a technique called “clean-piping” or “clean-scanning” (Oliva 2012). Temporary adjustments and using substitute vehicles violate state laws and constitute lenient regulatory monitoring.

One might also wonder about attempts to fraudulently fail a vehicle in order to charge the customer for making unnecessary repairs. Such attempts are on average both more difficult and less profitable than passing the vehicle. The difficulty of fraudulent over-stringency is that it involves deceiving both the state and the customer, who in this case have aligned incentives—customers want to avoid expensive repairs while the regulatory agency wants to ensure proper testing. With each facility in

our sample having an average of 58 competitors within a two-mile radius, most customers can easily retest their “failed” vehicle at another facility. The facility attempting such fraud risks losing the customer’s future business for testing and cross-selling. Furthermore, a customer who believes that his or her vehicle was falsely failed can verify this at another facility and can easily sully the fraudulent facility’s reputation on consumer websites such as Yelp, which can result in lost sales (Luca 2011). Such customers can also file a complaint with the regulatory agency, which increases the likelihood of a state investigation. Finally, the incentives for fraudulent failure are weak, even for facilities that might cross-sell repairs to remediate the problem. Emissions repair bills are limited to the \$450 necessary to receive a one-year emissions waiver. This one-time repair bill is worth considerably less than the average annual service and repair bill the facility could charge in the following year. Edmunds.com, for example, estimates the annual service and repair costs of a five-year-old Chevrolet TrailBlazer at \$2,089, with older vehicles having even higher cross-selling potential.

A facility extending leniency risks invoking a state investigation that can lead to the suspension or revocation of its monitoring license, as well as penalties that, in New York, can reach \$15,000 for the first offense and as much as \$22,500 for each subsequent offense (Navarro 2010). In addition, a facility found to be engaging in fraud by extending leniency risks being reported by the media (for an example, see States News Service 2009a), which can sully its reputation. Investigations can take the form of an undercover investigator bringing in a vehicle known to have excessive emissions. Such covert investigations are typically triggered by the regulatory agency observing suspicious patterns of test results. Because one agency regulates all the facilities in a given state, reputations for leniency that regulators associate with particular brands or subsidiaries are likely to be both salient and long-lasting. Reputational spillovers across states are less likely, although regional agency cooperation (through the U.S. Environmental Protection Agency, for example) could facilitate the transfer of information on likely offenders.

Organizational Scope and Monitoring Leniency

The vehicle emission testing market in New York consists of thousands of private-sector inspection facilities, with substantial variation in their organizational scope and private governance. Emission testing is a minor source of income for licensed facilities in New York, due to the state-mandated price of approximately \$20. All testing facilities in our sample are multi-product/service establishments—gasoline retailers, car dealers, or service and repair shops—for which inspections are a secondary business line. Each of these three types of business has different incentives and disincentives for leniency, based on profitable opportunities to cross-sell to repeat customers (Hypothesis 1a) and on the moral hazard risk associated with those cross-selling opportunities (Hypothesis 1b). We present these three types of business in Figure 1.

Gasoline Retailers: Gasoline retailers are unlikely to benefit from providing lenient monitoring because the cross-selling opportunity evokes very low profit opportunities. Gasoline retailing consists of small, low-margin gasoline sales transactions involving little customer loyalty. Retail gasoline is highly competitive, with publicly posted prices as the biggest drivers of consumer choice. A 2009 survey showed that price was the primary factor in gas station choice for 70% of consumers, with 59% willing to drive five minutes out of their way to save five cents per gallon (National Association of Convenience Stores 2009). The average gasoline retailer earns only \$0.02-\$0.03 per gallon in pre-tax profit and is therefore reliant on convenience store sales and sales of related products (National Association of Convenience Stores 2009). The upside of leniency is thus quite limited for gasoline retailers. Gasoline retailers also face some potential downside from exhibiting lenience, in the form of the risk of a poor reputation spilling over to their primary business, which relies on customers trusting that gasoline is unadulterated and precisely measured (Olmstead and Rhode 1985). Yet because government agencies regularly inspect fuel and pumps and because the magnitude of adulteration is limited by engines' combustion requirements (above which engine malfunctions would trigger customer complaints), we expect customer fear of moral hazard from gasoline retailers to be relatively low. This lack of moral hazard risk, however, cannot compensate for the very low profit opportunities associated with cross-selling gasoline and

therefore has little impact in motivating leniency. Figure 1 illustrates that customers perceive low risk of moral hazard from gasoline retailers, but that gasoline retailers face little profit opportunity from providing lenient monitoring.

Service and Repair Facilities: For service and repair facilities, however, the opportunity to cross-sell products and services less price-sensitive than gasoline is a strong incentive to provide leniency in emissions testing. Because vehicles with emissions problems tend to have other mechanical problems needing large and frequent repairs, mechanics have strong incentives to keep these cars on the road. Annual car repair expenditures average \$600 to \$800 and five- to ten-year-old vehicles are likely to require more than double this annual amount.³ Gross margins on repair services average around 50% (First Research 2010)—much greater than the margins on emissions tests and vastly superior to the small margins on gasoline sales. Together, these frequent, high-margin repairs generate a profitable opportunity for service and repair shops to maintain long-term customer loyalty.

While, at first glance, the profitability of repairs might suggest these facilities would benefit from failing cars, it is important to consider that the upside of an immediate repair is limited by state regulations, which cap necessary emission-related repairs at \$450. If a vehicle owner spends \$450 on repairs and her vehicle still fails, she can receive a one-year waiver allowing her to keep using it. As demonstrated in the Appendix, many of these customers will not return to the same facility the following year, seeking a more lenient facility instead. Although stringency does limit the already low likelihood of detection and punishment, a facility that chooses stringency is valuing a moderate one-time payment more than a considerably more lucrative stream of future service and repair work.⁴ Since all vehicles require regularly scheduled maintenance and face unexpected mechanical failures, long-term relationships with

³ These *Household Spending* data from New Strategist (2007) also include very new models and thus underestimate expenditures on the older cars most likely to need leniency. Values from Edmunds.com, as noted earlier, are much larger and more representative of older cars.

⁴ Annual maintenance costs for relatively recent vehicle models can be calculated at <http://www.edmunds.com/tco.html>.

repeated transactions are highly valuable. All these factors create strong incentives for service and repair facilities to provide leniency.

Yet leniency also presents moderate risks due to potential moral hazard in the cross-sold repairs. Customers may fear that a firm willing to cheat the state might also deceive its own customers about repair services. The service and repair of existing problems are experience goods and thus present limited moral hazard risk in long-term relationships due to the ex-post verifiability of repair quality (Rey and Salanie 1990). Unnecessary repairs, however, may be credence goods if the vehicle is asymptomatic and the customer does not seek a second opinion. Customers therefore bear some risk that service and repair facilities will engage in moral hazard (Taylor 1995; Pesendorfer and Walinsky 2003; Schneider 2011).

For customers, the degree of moral hazard risk depends largely on the existence of reputational mechanisms and long-term customer relationships. When customers have long-term or potentially long-term relationships involving multiple transactions with repair/service facilities, or when reputation is observable (Klein and Leffler 1981), this moral hazard risk is limited. If there is any likelihood that customers will detect unnecessary repairs and therefore take their future business elsewhere, the repair facility will be much less likely to engage in moral hazard.

But moral hazard is a serious risk for customers without long-term relationships, especially for those unlikely to return, in which case the facility would have an incentive to maximize profits from the one visit. As in a one-shot trust game without punishment (Kreps 1990), the profit-maximizing behavior for such facilities may be to be stringent, hoping for immediate repair business to remediate an emissions problem. The same may be true in cases of facilities facing extreme financial distress, the equivalent of a high discount rate in a trust game, which would also make stringency in hopes of immediate repairs more desirable. Yet even in these cases, a savvy customer retains the right to retest his or her vehicle elsewhere, so strategic stringency in generating repair business is of somewhat limited efficacy. Thus, while there may be some conditions under which service and repair facilities have incentives for stringency, we expect the potential for long-term high-margin repeat business to promote leniency on average. We represent this likely high level of leniency in Figure 1.

Car Dealerships: Car dealers also enjoy profitable cross-selling opportunities from leniency. Dealers are likely to garner customer loyalty, which can generate hundreds or even thousands of dollars in profits if a customer returns to purchase a vehicle, even if that purchase is several years later. A Bain & Company survey of 1,800 car dealership customers found that those receiving quality service at a dealership are much more likely to purchase their next car there (Lamure, Hoffmann, and Flees 2009). In addition, most car dealerships also engage in service and repair activities and thus have further incentives to retain customers through lenient emissions testing. Like the service and repair facilities discussed above, dealerships also face the risk of reputational spillover from leniency due to moral hazard risk, but this is more likely to be the case for smaller used car dealerships selling older cars.⁵ Buyers of new vehicles face low risk of moral hazard because the aesthetic and performance attributes that define new vehicle quality are readily observable and widely documented. Furthermore, both new and late-model used vehicles are protected by long warranties (Spence 1977).⁶ With little risk of moral hazard in the cross-sold market, the risk of reputational spillovers is unlikely to reduce leniency on the part of car dealerships. Consequently, as we show in Figure 1, we expect that car dealerships' high profit opportunities and low moral hazard risk lead to high levels of testing leniency.

Consistent with the hypothesized leniency in Figure 1, many states have well understood the conflict of interest between emissions testing and service, repair, and sales activities, leading them to implement "test-only" facilities. The U.S. Environmental Protection Agency's early studies found that using test-only facilities reduced emissions by twice as much as when testing facilities were also allowed to perform repairs (Cohn 1992). When the Wisconsin legislature designed its emissions testing program in 1979, it was so concerned about these potential incentive problems that it prohibited inspection facilities from being "engaged in the business of selling, maintaining or repairing motor vehicles or of

⁵ Among emissions tests conducted by car dealers, the vast majority (78%) were conducted by dealers that sold new vehicles (some of which also sold used vehicles); only 22% were conducted by dealers that only sold used vehicles. As a robustness test described below, we also estimated distinct effects from these two types of car dealer.

⁶ Considerable increases in vehicle quality over the last decade have led nearly all manufacturers to certify (warranty) "pre-owned" vehicles sold at their dealerships, applying extensive warranties that extend up to 100,000 miles (Sultan 2008).

selling motor vehicle replacement or repair parts” (Franzen 2008: 8). This suggests that car dealers and service and repair facilities profit much more than gasoline retailers from lenient monitoring, as represented in Figure 1. Hence, in our empirical analysis that tests Hypotheses 1a and 1b, we compare gasoline retailers to these two other facility types.

Private Governance and the Stringency of Monitoring

Vehicle owners across New York State choose from thousands of private inspection facilities, all licensed by the state to conduct emissions tests but with substantial variation in monitoring leniency and private governance. These governance structures include corporate-owned subsidiaries, branded franchises, and independent establishments. Compared to independent establishments, we expect subsidiaries and branded facilities to be more stringent, as expressed in the reputation-based Hypotheses 2a and 2b. We also expect, as expressed in Hypothesis 3, that subsidiary status will reduce leniency more than branded affiliation.

Data and Measures

Our primary dataset, obtained from the New York State Department of Environmental Conservation, contains all dynamometer vehicle inspections conducted from 2000 through 2004 on gasoline-powered vehicles weighing less than 8,500 pounds at service and repair facilities, gasoline retailers, and car dealers. We linked these facilities via name and address to Dun & Bradstreet (D&B) data, obtained from the National Establishment Time-Series (NETS) database, to obtain a primary Standard Industrial Classification (SIC) code and unique D&B identifier (DUNS number) for each facility and its ultimate parent organization.

Vehicle and test characteristics. We identify each vehicle’s Vehicle Identification Number (VIN), make, model, year, weight, odometer reading, inspection date, and whether or not it passed. We identify specific vehicle models by creating *vehicle model* fixed effects for each unique combination of the first eight digits of a VIN, which identify the vehicle’s manufacturer, year, model, body, and engine specifications. Since the inspections are tests of tailpipe exhaust, a vehicle passes only if it scores below a

government-mandated threshold for all three constituents: hydrocarbons (HC), carbon monoxide (CO), and nitrogen oxides (NO_x). Our data do not contain any information about the vehicle owners.

Organizational governance. We created three measures of organizational governance. We consider an inspection facility to be a *branded facility* if its name includes a brand name associated with a gasoline retailer (e.g., Shell, Mobil), a service/repair chain (e.g., Bridgestone, Midas, Meineke), or a vehicle make (e.g., Ford, BMW). We consider an inspection facility to be a *subsidiary facility* in a particular year if its DUNS number differs from that of its headquarters or if it shared its headquarters' DUNS number with another facility that year. We considered a facility to be an *independent facility* if it is neither branded nor a subsidiary. Because a facility's name, primary SIC code, and headquarters can change over time, we coded these variables at the facility-year level. Tallies of unique facility-years and inspections associated with independent, branded, and subsidiary facilities in our sample are reported in Table 1. The branded and subsidiary designations are not mutually exclusive. Our sample has 663 branded subsidiaries, such as Bridgestone, Firestone, Goodyear Auto Service Centers, and Pep Boys facilities.

INSERT TABLE 1 HERE

Because the incentive to protect a brand may increase with the number of branded facilities, we created *branded siblings* by logging the number of facilities in our sample that shared a brand (after adding one). Similarly, we created *subsidiary siblings* as the logged number of facilities in our sample that shared a parent company. Our results are robust to measuring these constructs as raw counts (without taking logs).

Organizational scope. We categorized inspection facilities each year into one of three mutually exclusive industries based on the primary three-digit SIC code assigned that year by D&B. As described below, we used facility name to categorize facility-years that failed to match D&B data or that were missing SIC codes. We created a dichotomous variable, *car dealer*, coded “1” when a facility's primary three-digit SIC code was 551 (“Motor vehicle dealers—new and used”) or 552 (“Motor vehicle dealers—

used only”) or—if we did not know its SIC code because we could not match the facility to D&B data—when the facility name included a term that implied a car dealership, such as “sales,” “auto mall,” “used car,” or a vehicle brand name (e.g., “Acura”). The full list of search terms used to identify car dealers is available from the authors. *Car dealer* was coded “0” when the facility has a different primary SIC code or—when we lacked an SIC code—when its facility name did not include the aforementioned terms indicating a dealership. Of the 1,623 unique facility-years we associated with car dealers, 1,291 (80%) were classified based on SIC codes and the rest based on facility names. Tallies of unique facility-years and inspections associated with gasoline retailers and service and repair facilities are presented in Table 1.

We created a dichotomous variable, *gasoline retailer*, to denote inspection facilities primarily engaged in selling gasoline. These were mainly facilities with a primary three-digit SIC code of 554 (“Gasoline service stations”). For inspection facilities we could not match to D&B data, we identified as gasoline retailers those with company names including any of the following terms: Amoco, Arco, BP, Chevron, CITGO, Esso, Exxon, Getty, Gulf, Marathon, Mobil, Phillips, Shell, Sunoco, and Texaco. Of the 2,998 unique facility-years in our sample that were associated with gasoline retailers, 2,829 (94%) were classified based on SIC codes and the remainder were classified based on these company names.

We created a dichotomous variable, *service and repair facility*, to denote facilities of which the primary activity in a given year was conducting vehicle service and repairs or selling parts; for simplicity, we refer to these simply as service and repair facilities. We identified these facilities as those for which the primary three-digit SIC code in a given year was 553 (“Auto and home supply stores”), 753 (“Automotive repair shops”), or 769 (“Miscellaneous repair shops and related services”). For facility-years to which we could not match D&B data and that were not already categorized as a gasoline retailer or car dealer, we identified as service and repair facilities those that reported repair data (in addition to emissions data) to the state regulatory agency.⁷ Of the 9,680 unique facility-years associated with service

⁷ Although the state requests repair data from inspection facilities, the data are incomplete and self-reported. Thus, although these repair data enable us to identify some facilities that conducted repairs, they do not provide a comprehensive list of repair facilities nor even of the repairs conducted at the reporting facilities.

and repair facilities, 7,464 (77%) were classified based on SIC codes and the remainder were classified based on reporting repair data to the state agency.

It is important to note that many of the facilities in our sample likely engage in multiple activities. Many car dealers and gas stations do some service and repairs, while some service stations may also sell gasoline. This measurement error in our scope variables makes our identification more difficult and, if anything, biases against us finding results. If some gas stations or car dealers partially behave like repair and service facilities, identifying differences between these categories is even more empirically difficult. We would therefore expect the true differences in leniency between firms of different scope to be stronger than we empirically identify.

Additional facility/market characteristics. We measured *facility inspection volume* as the log of the average number of monthly inspections a facility conducted during the two months preceding a focal inspection. We measured inspection *competition* by logging (after adding 1 to accommodate zero values) the number of other gasoline retailers, car dealers, and service and repair facilities that conducted inspections within the focal facility's five-digit ZIP code.⁸ We controlled for neighborhood wealth by logging the median household income for the focal facility's geographic area (Census Place), based on 2000 U.S. Census data (*facility neighborhood's median household income*).

We also control for whether an inspection facility is a member of the AAA Approved Auto Repair Network, operated by the American Automobile Association (AAA). AAA certifies repair facilities after an AAA specialist “inspects the facility for cleanliness, proper tools, adequate technical training, and appropriate technician certifications”; confirms that at least 90% of the facility's customers are satisfied with their repair work; and “checks the facility's reputation with government and consumer agencies” (American Automobile Association 2010). The AAA certification is akin to third-party certification processes in other industries, which are used to signal honesty and convince customers and governments that adopters have implemented world-class management practices governing elements such

⁸ For a more detailed examination of the impact of competition on leniency, see Bennett et al. (2012).

as labor, quality, and environmental affairs (e.g., Corbett, Montes-Sancho, and Kirsch 2005; Darnall and Sides 2008; King, Lenox, and Terlaak 2005; Terlaak and King 2006). AAA certification may be associated with stringency either through selection processes at the facility level (stringent firms seek certification) or customer level (law-abiding customers seek AAA facilities) or through treatment effects on the facility (monitoring by AAA reduces fraudulent leniency). We coded *AAA-certified facility* “1” when a facility was a member of the AAA Approved Auto Repair Network (and “0” otherwise) based on data obtained from the websites of New York State’s various AAA clubs and by calling the clubs. Forty-four facilities (209 unique facility-years) in our sample are AAA-certified.

Summary statistics and correlations for our primary sample are reported in Table 2, Panel A. In our sample, 92% of vehicle emissions tests were passed; only 8% were failed. For our mutually exclusive activity scope categorizations, 70% of emissions tests were conducted at service and repair facilities, 23% at gasoline retailers, and the remaining 7% at car dealers. 79% of emissions tests in our sample were conducted by independent facilities, 16% by branded facilities, and 10% by subsidiaries.⁹ On average, facilities in our sample conducted 86 vehicle emissions tests per month. Table 2, Panel B presents descriptive statistics by scope and governance designation. Vehicle characteristics are relatively consistent across these categories, although car dealers have the youngest, lowest-mileage cars while service/repair stations have the oldest, highest-mileage cars. Consistent with this, car dealers have the highest pass rate, while service/repair stations have the lowest. Table 2, Panel C presents the pairwise correlations for our main explanatory and control variables.¹⁰

 INSERT TABLE 2 HERE

⁹ The sum of these figures exceeds 100% because some facilities are both branded and subsidiaries, as shown in Table 1.

¹⁰ Because gas retailers, service/repair stations, and dealerships are mutually exclusive categorical variables, the negative correlation between these variables is purely a function of their relative frequency. Whereas two mutually exclusive categories would be perfectly negatively correlated at -1.0, with three categories, two are always coded 0 whenever the other is coded 1, which increases their correlation from the -1.0 baseline. Consequently, when considering the correlations to gasoline retailers being -0.83 for service/repair stations and 0.15 for car dealers, the more negative correlation of -0.83 simply reflects the greater number of tests at repair facilities than at car dealers.

Empirical Approach and Results

Preliminary Analysis: Customer Loyalty as an Incentive for Leniency

The arguments supporting Hypothesis 1a require a demonstration that customer loyalty is an incentive for leniency. While this was observed in Hubbard's (2002) study of emissions testing within a local market in California and is supported by interviews with regulators and by government documents (California Bureau of Automotive Repair 2011), we verified this incentive for customer loyalty using our larger sample from a different state. We used logistic regression to estimate the probability that a customer would return to a facility as a function of whether his or her vehicle failed its prior test there, controlling for vehicle and facility characteristics. Our approach, detailed in the Appendix, is akin to Hubbard's (2002), but we rely on five years of data—millions of inspections from several thousand facilities as opposed to Hubbard's 29 facilities—reducing the risk that our results are idiosyncratic to a limited number of firms. Our much larger sample also allows us to better control for non-independence in the error structure by clustering at the facility level.¹¹ Our results indicate that the probability of a customer returning to an inspection facility where his or her vehicle had previously been inspected declined by 9.8 percentage points when the vehicle failed (Column 1 of Table A-1), an 18% reduction from the sample average return rate of 53%. Failing an emissions test can be costly for the vehicle's owner, who may need to have it repaired or to sell it to someone in a state with less stringent emissions requirements. An inspection facility that fails a vehicle risks losing that customer not only for future emissions testing but also for its primary business activity.

Estimating the Stringency of Monitoring

Having demonstrated that customer loyalty can be an incentive for leniency, we now describe our approach to empirically testing our hypotheses on how organizational scope and governance affect a

¹¹ As a robustness test, we clustered errors to account for relationships between facilities. Specifically, we clustered our error terms at the headquarters level for all subsidiaries, at the brand level for all branded facilities (except branded subsidiaries, which were clustered by headquarters), and at the facility level for all independent facilities. The results were virtually identical.

firm's leniency. In doing so, we attempt to control for many other factors that might also affect the likelihood of passing a vehicle, including test time and location and vehicle-specific factors. We then interpret the higher average pass rate associated with a particular type of facility as an indication of leniency, an approach used in previous studies of vehicle emissions testing (Gino and Pierce 2010; Pierce and Snyder 2008) and based on well-established measures of risk-adjusted performance in the healthcare productivity literature (e.g., Cutler, Huckman, and Kolstad 2010; Huckman and Pisano 2006). The risk of omitted-variable bias associated with this technique is substantially mitigated by our detailed vehicle data and panel structure. We use the following model, in which the unit of analysis is the individual vehicle emissions test, to estimate the probability that a vehicle passes an emissions test:

$$(1) \text{Pass}_{ijt} = F(\text{Governance}_{it}, \text{Scope}_{it}, \text{VehicleCtrls}_{jt}, \text{Competition}_{it}, \text{TestCtrls}_{it}, \text{FacilityCtrls}_{it})$$

where $F(\cdot)$ is the logit function. Pass_{ijt} is a dummy coded "1" if vehicle j passed its inspection at facility i on date t and coded "0" if it failed. Governance_{it} represents our two variables that log the number of facilities that share the focal facility's brand or parent company.¹² Scope_{it} represents a series of dummy variables that indicate whether the facility is a gasoline retailer, car dealer, or service and repair facility in year t (gasoline retailer is the omitted category). VehicleCtrls_{jt} includes characteristics of vehicle j inspected in year t known to affect a vehicle's likelihood of passing an emissions test (National Research Council et al. 2001: 237). These factors include vehicle model fixed effects based on the first eight digits of the vehicle's VIN. We include the vehicle's odometer reading to control for deterioration from usage. We include odometer as its level, squared, and cubed values since we have no priors about the specific functional relationship between vehicle usage and pass rates and wish to allow for flexibility in the functional form. We include a full set of dummies to control for vehicle age (in years) at the time of the test.¹³

¹² Models using dummy variables for brand and subsidiary produce substantively identical results.

¹³ We note that while the length of customer relationship is theoretically interesting, we are unable to include it in our model because it is endogenously determined by the prior leniency a customer experienced from the facility.

FacilityCtrls_i includes facility characteristics that might influence pass rates, including dummy variables denoting the first three digits of the facility ZIP code to control for geography-based differences between facilities (e.g., climate, population density), the *facility neighborhood's median household income*, the facility's *inspection volume*, and a dummy variable that indicates whether or not the facility was AAA-certified. Dummies for three-digit ZIP codes and inspection month in *TestCtrls_{it}* account for the influence of ambient conditions on vehicle emissions (National Research Council et al. 2001: 238). Because emissions test standards changed in the focal state during 2003, we split the 2003 year dummy into two dummies to distinguish the periods before and after the change. Because research suggests that greater competition can affect quality (Banker, Khosla, and Sinha 1998; Becker and Milbourn 2011), we include *competition_{it}*, which incorporates the number of other inspection facilities within the same five-digit ZIP code.

Baseline model: Vehicle model fixed effects. We used logistic regression to estimate the likelihood that a vehicle passed its emissions inspection. In our baseline model, we include unconditional vehicle-model fixed effects based on the first eight digits of the VIN to control for differences in pass rates between vehicle models. Studies have shown that bias is negligible when unconditional fixed-effects logit models have at least 16 observations within each group (Katz 2001; Coupé 2005; Greene 2004). We pursue a conservative approach by limiting our sample to vehicle models with at least 100 inspections and at least five emissions tests that failed. These restrictions facilitate model convergence and ensure that the fixed effects do not result in biased estimates.

Models 1 and 2 in Table 3 enter our scope and governance variables separately into different regressions. Model 3 represents our fully specified model; we report coefficients and average marginal effects (Column 3). In each of these models, standard errors are clustered by facility. Our results are virtually identical when we cluster by firm/brand and our estimates are even more precise when we cluster by the particular vehicle (VIN) or by vehicle make. The results indicate that car dealers are substantially more lenient than gasoline retailers (the omitted firm-scope category), which supports Hypotheses 1a and 1b. The average marginal effect indicates that car dealers are 2.0 percentage points

more likely to pass the same vehicle. Given the sample average fail rate of 8%, car dealers are approximately 25% less likely to fail a vehicle (calculated as $0.020 \div 0.08$). The positive coefficient on service and repair facilities is consistent with our hypothesis of leniency, but it is not statistically significant.¹⁴ As predicted by Hypotheses 2a and 2b, our results indicate that subsidiary and branded facilities are less lenient than independent facilities, the omitted governance category. Specifically, the average marginal effects reported in Column 3 indicate that, compared to independent facilities, an additional facility that shares the focal facility's brand is associated with a 0.2-percentage-point decline in the probability of passing a given vehicle, which corresponds to a 2.5% increase in the failure rate from the 8% sample average failure rate. Each additional facility sharing the focal facility's parent company is associated with a 0.9-percentage-point decline in the probability of passing a vehicle, an 11% increase in the failure rate from the 8% sample average failure rate. This difference, whereby an additional subsidiary deters leniency more than an additional facility that shares a brand, is statistically significant ($p < 0.01$ from a post-estimation Wald test), which supports Hypothesis 3.¹⁵

 INSERT TABLE 3 HERE

Vehicle fixed-effects models. While our baseline model controls for many characteristics of vehicles, inspection facilities, and testing conditions, it remains possible that omitted variables might be

¹⁴ Estimating Model 3 without clustering, as in Hubbard (1998), yields much smaller standard errors, such that one might conclude that the effect of service and repair stations is highly statistically significant ($p < 0.001$). The effect of service and repair stations is also highly statistically significant when we tried clustering standard errors on vehicle (VIN) or vehicle make ($p < 0.01$ in both cases). Our results suggest that previous work without clustering might suffer from Type I error. We also note that our coefficient may be biased downward and less precise due to a subset of service/repair facilities that fraudulently fail naïve customers to earn immediate repair business. While these cases may be relatively uncommon (and unidentifiable in our data), they may contribute to larger standard errors.

¹⁵ For those concerned that our large sample size warrants more stringent criteria to test hypotheses, we applied a more conservative variation of Leamer's formula that was originally developed for an OLS context with independent observations and constant variance (Leamer 1978: 114). We made conservative adjustments to accommodate our clustered context by using the number of clusters ($T=3530$ facilities) instead of the number of observations (n inspections) to calculate Leamer's Critical F and Critical z values as follows: Critical F = $T \cdot (T^{1/T} - 1) = 8.18$ and thus Critical $z = \sqrt{\text{Critical F}} = 2.86$. After estimating our model using OLS regression and clustering standard errors by facilities, the three hypothesized results found to be significant at conventional levels (*dealers*, *subsidiary siblings*, *branded siblings*) had t-values that also exceed Leamer's Critical z value of 2.86. Testing Hypothesis 3 via a Wald test confirmed that the magnitude of the subsidiary effect exceeds that of brandedness (Wald F=15.19, $p < 0.01$) and this F value exceeds the Critical F value of 8.18.

correlated with our key independent and dependent variables, thereby biasing our results. For example, if owners took only the worst of each vehicle model (e.g., 2001 Honda Civic) to gasoline retailers, this could potentially explain lower pass rates at these facilities. To control for time-invariant characteristics of each individual vehicle (e.g., its particular feature set and the conditions under which it was manufactured) that might lead to omitted-variable bias, we include conditional VIN fixed effects in Model 4. To further control for aspects related to a particular vehicle owner, which could result in changes in a vehicle's (unobserved) maintenance level and preferred type of inspection facility, Model 5 includes fixed effects for each VIN-owner pair. Because inspection stickers in our focal state are granted for one year but vehicles must be re-inspected within two weeks of a vehicle sale, we followed Hubbard (2002) in identifying ownership changes by a vehicle's inspection occurring in a calendar month different from that of its previous inspection. Thus, in Model 5 we include fixed effects denoting vehicle-owner pairs based on each unique combination of VIN and its inspection calendar month.¹⁶

We estimate both of these models with conditional fixed-effects logistic regression.¹⁷ Model 4 is identified only for vehicles that pass and fail at least once and Model 5 is identified only for vehicles that pass and fail at least once in the same calendar month. In both models, the coefficients on the firm-scope variables (service/repair stations and car dealers) and governance variables (subsidiary, branded, and AAA-certified) are identified only by those vehicles that switch between facility types at least once during our sample period. We therefore refer to these as “switcher models.” The key trade-off with these models, in comparison to the baseline model, is their superior control for unobservable factors, but at the expense of reduced sample sizes due to conditional logistic models dropping observations associated with vehicles (Model 4) or vehicle-owner pairs (Model 5) that lack variation in their inspection outcome.

¹⁶ This approach, for example, considers all April inspections of a particular vehicle to be associated with one owner, and all October inspections of that vehicle to be associated with a different owner. We consider these to be two distinct vehicle-owner pairs. If an owner really had tested his or her vehicle in different months, this would simply create two fixed effects for that vehicle-owner pair, which does not bias our estimates.

¹⁷ Given how few inspections (observations) there are for each vehicle (VIN), when estimating these within-vehicle models, unconditional fixed effects risk producing biased estimates and thus we rely on conditional fixed effects.

Furthermore, we are unable to cluster errors at the facility level because it is not nested within our conditional fixed effects, so we clustered at the vehicle (VIN) level.¹⁸

The results of Models 4 and 5 (Columns 4 and 5 of Table 3) support all of our hypotheses. The positive statistically significant coefficients on service and repair stations and car dealers continue to provide evidence that these facility types offer greater leniency than gasoline retailers do, lending additional support for Hypotheses 1a and 1b. Leniency declines for facilities that share a brand or a company parent, supporting Hypotheses 2a and 2b. The point estimates continue to suggest that the deterrence on leniency is stronger for subsidiaries than for brands; the magnitude of the negative coefficient on subsidiaries exceeds that of the negative coefficient on brands and the difference is statistically significant ($p < 0.01$ from post-estimation Wald tests from both of these models). This supports Hypothesis 3. While our results from these vehicle conditional fixed-effect models are generally consistent with our baseline vehicle model fixed-effects results, we highlight two distinctions. First, Models 4 and 5 yield substantively larger effects for *service and repair facility*, *subsidiary*, and *branded facility* and smaller effects for *dealer*.¹⁹ Second, Models 4 and 5 yield coefficients on *AAA-certified facility* that are statistically significant and negative, indicating considerably less leniency at these facilities than at noncertified facilities.

In summary, our conditional fixed-effects logistic models—despite being identified for considerably fewer observations than our baseline model—appear to more precisely identify the role of scope and governance because they better control for omitted vehicle-level factors. One potential explanation is that the worst vehicles, which are frequently repaired, are being tested at repair facilities, thereby biasing downward the parameter estimates in our baseline models. Similarly, the increased

¹⁸ As described below, our results were nearly identical when we estimated this model using OLS with fixed effects, clustering standard errors either by facility or by vehicle.

¹⁹ Estimating the Model 3's specification on the smaller sample used in Model 4 yielded coefficients very similar to those of Model 3 when estimated on its larger sample. This rules out the change in sample as the cause of the increased effects.

parameter estimate for AAA-certified facilities might be biased upward (toward zero) in the baseline models if the best-maintained vehicles were tested at these locations.

We must be careful, however, in drawing these conclusions, as the error corrections in these models were necessarily less conservative because the conditional fixed-effects logistic regression models could not be clustered at the more conservative facility level.²⁰ While we pursued a second-best approach—clustering at the vehicle level—our results might nonetheless suffer from Type I error if the standard errors are unduly small. To assess this, we reestimated these specifications (and identical samples) as linear probability models (OLS), which enabled us to cluster by facility. Like our logit results clustered by vehicle, these OLS results clustered by facility continued to yield statistically significant positive coefficients on the service/repair facility and car dealer variables and statistically significant negative coefficients on the subsidiary and branded variables. Furthermore, the magnitude of the subsidiary effect continues to significantly exceed that of the brandedness effect ($p < 0.01$ from post-estimation Wald tests). These results bolster our confidence that the significant results yielded by our conditional fixed-effects logit models, reported as Models 4 and 5, are not attributable to Type 1 error that might be associated with our inability to cluster by facility.

Robustness Tests

One potential concern about our results is that they might be vulnerable to omitted-variable bias if vehicle owners select facility types such that unobserved vehicle attributes are correlated with facility type. For example, vehicle owners with poorly maintained vehicles might seek independent facilities that lack the private governance mechanisms associated with branded and subsidiary stations. Such sorting would result in vehicles in poor condition being disproportionately tested at independents, which would lower the pass rates for that type of facility. Because we hypothesize higher pass rates at independents than at branded or subsidiary stations, such sorting would constitute a bias against our hypothesized results. But what if vehicles in poor condition were more likely to be inspected at service and repair shops

²⁰ In conditional fixed-effects logistic regression models, standard errors can only be clustered at levels nested within the conditional fixed effect.

and car dealers than at gasoline retailers, so that if they needed repairs, they could be done at the same time? This would result in cars with a higher probability of failing being tested at service and repair shops and car dealers and would depress pass rates at these facility types. Since we hypothesize that such facilities have higher pass rates than gasoline retailers, this type of sorting would also result in bias against our hypothesized results.

The potential for this form of sorting to result in omitted-variable bias (in favor of or against our hypothesized result) is already limited to potential time-variant effects (i.e., vehicle owners changing preferences over time), given that our model with VIN-month fixed-effects (Column 5 in Table 3) controls for time-invariant vehicle owner preferences for facility type (i.e., time-invariant preferences of vehicle-owner pairs). Nevertheless, we sought to assess whether our results might be influenced by vehicle owners self-selecting into different facility types if their vehicles were especially likely to need repairs. We therefore reestimated our primary model (reported in Column 3 in Table 3) on two subsamples of vehicles with particularly high risk of failing the inspection. The first includes vehicles that have failed at least once in the prior three years but are still on the road. These cars are, per our investigation, substantially more likely than the general population of cars to fail again. Because many of these cars might have been repaired, we use a second subsample: vehicles that almost failed in the previous year. We build this “at risk” sample by calculating the average annual deterioration (increase in emissions) of passing cars for each make/model group (8-digit VIN). Any car with results less than this amount above the failing threshold would be more than 50% likely to fail in the next year under normal deterioration rates. Estimated on both of these samples of vulnerable vehicles, our model yielded results (Columns 1 and 2 of Table 4) that are very similar to those of our main sample, despite substantially smaller observation numbers. Our results were nearly identical when estimated using OLS with fixed effects, clustering standard errors either by facility or by vehicle.

INSERT TABLE 4 HERE

Another potential concern with our analysis is that, compared to independent facilities, subsidiaries and branded companies might have access to greater financial and legal resources with which to defend themselves against government scrutiny and might therefore feel less threatened by government

sanctions and be more prone to lenience. This would constitute a bias against our hypothesized results, given that we predict that subsidiaries and branded facilities will be less lenient.

One might also be concerned that, compared to independent facilities, subsidiaries and branded companies' superior access to financial and legal resources might lead politicians (and by extension, regulators) to extend them preferential treatment. A firm's financial resources might serve as an effective deterrent to prosecution, especially in settings in which the relevant regulatory agencies are underfunded or highly vulnerable to political pressure. However, we do not believe this to be a significant concern in our context because prosecutions were conducted by the New York State Attorney General Eliot Spitzer, who was well known for aggressively prosecuting corporate offenders.

Independent, subsidiary, and branded facilities might also differ in terms of financial stress, which could affect the extent to which they might balance short-term versus long-term profitability. To assess this, we examined whether financial stress, measured by Dun & Bradstreet's PAYDEX scores, substantially differed between independent facilities and either branded or subsidiary facilities. PAYDEX is an "indicator of how a firm paid its bills over the past year... from 1 to 100, with higher scores indicating better payment performance" (Dun & Bradstreet 2012). Average financial stress levels were nearly identical across all three groups: independents averaged 68.2, branded facilities 68.3, and subsidiaries 67.6. T-tests revealed that these differences were not statistically significant at the 10% level. T-tests comparing independents to the other types of facility indicated that their differences were not statistically significant at the 10% level. These comparisons suggest that, despite differences in legal structure, the financial stress of independents was indistinguishable from that of subsidiaries and branded facilities.

One might also be concerned that inspections might constitute a smaller portion of total business for gas retailers than for car dealers and service and repair stations and that gas retailers therefore made less effort in passing vehicles or attracted less qualified or experienced technicians. This concern is attenuated by two facts in our empirical context. First, as noted above, all technicians who conduct emissions tests in New York State are state-certified, which ensures a minimum level of knowledge, experience, and education (as described earlier). Second, gas stations in our dataset actually conduct slightly more inspections per day (2.47) than do service/repair facilities and car dealers (2.41). This suggests that inspection experience is likely to be quite similar across these facility types.

One might be concerned that our results are driven by differences across facility types in their expected relationship durations rather than by our hypothesized mechanisms. In particular, might the most lenient types of facility—independents, service and repair shops, and car dealers—expect longer-term relationships than other types of facility and therefore have the most to gain by falsely passing vehicles? To assess this, we calculated the average duration of emissions testing relationships. In fact, the average relationship at service/repair stations and car dealers (1.61 years) was slightly shorter than that at gas retailers (1.70 years), which would present a bias against our result. The average relationship at subsidiaries and branded facilities (1.58 years) was only slightly shorter than that at independent facilities (1.65 years), a disparity most likely generated by the increased leniency of the independents.²¹

A potential concern with our Dun & Bradstreet industry designations is that some gasoline retailers may also be engaged in service and repair. Similarly, some service/repair facilities may sell gasoline. We believe that selling gasoline as a secondary service would only weaken a repair facility's motive to engender loyalty, as gasoline retailing is less dependent on customer loyalty than service/repair activities are. This possibility therefore serves as a bias against our hypothesized result. Similarly, a gasoline retailer providing service and repairs would have higher incentives for leniency, which would also bias against our results. Consequently, we expect that our estimates of the impact of scope on leniency are understated due to potential imprecision in classification.

As an extension, we sought to understand whether brandedness impacts different types of facility differently. To do so, we estimate on our full sample a model that includes brandedness interacted with the facility's scope and subsidiary status. The results, reported in Column 3 of Table 4, are very similar to those from our main models, except that the positive coefficient on *car dealer* is considerably smaller and the interaction between *branded siblings* and *car dealer* is positive and statistically significant. These results indicate that, while dealerships are on average more lenient than other facilities, branded dealerships are even more lenient, contradicting the hypothesized and average effect of brandedness on leniency. We expect this is because branded dealerships are almost exclusively new-car dealerships, which present much lower risk of moral hazard to customers than do used-car dealerships.

²¹ We calculated these values using only the youngest vehicles in our sample (those 5-7 years old) that had never failed, in order to avoid failure-induced switching. Despite limiting our sample for this test to those vehicles least likely to require leniency, some leniency is still likely occurring, which likely accounts for the minor differences in customer duration.

Discussion and Conclusion

In this study, we examine how private regulatory monitors balance market and institutional forces. Private monitors seek to avoid government sanctions for lenient enforcement, but such leniency can enhance customer loyalty and thus profitable cross-selling. Firms capable of cross-selling products and services must consider both the opportunities of profit gains from leniency and the risks that customer fears of moral hazard will lead to reputational spillovers across services. Firms operating under different governance structures face different costs of sanctions and different impacts on reputation, resulting in some firms being constrained by additional monitoring conducted by corporate parents and brand owners. Consequently, we hypothesized that product scope and governance mode influence the strategic choice of leniency. Our empirical evidence supports these hypotheses.

We observe more leniency from car dealerships than from gasoline retailers, revealing the importance of organizational scope and the potential conflict between the pursuit of customer loyalty and monitoring stringency. Dealers seek to cross-sell high-margin new vehicles to loyal customers and can promote customer satisfaction through emissions testing leniency. Circumstances are very different for gasoline retailers, however, which are unable to profit greatly from cross-selling gasoline and which risk distrust of measurement accuracy when known by customers to engage in fraudulent behavior. Gasoline retailers have little to gain and quite a bit to lose from helping noncompliant vehicles to pass inspection.

We also find some evidence that service and repair shops are more lenient than gasoline retailers. Compared to car dealers, the smaller magnitude of our leniency estimate for service and repair stations is consistent with our hypothesis that service and repair facilities suffer greater risk of reputational spillovers due to customer fear of moral hazard. Moreover, service and repair facilities serve some customers who are unlikely to return, which might create incentives for a facility to fraudulently fail those vehicles. Pooling these customers with long-term customers may lead to these countervailing incentives cancelling one another out, resulting in a smaller or unidentified pooled effect. Despite our inability to separate these customer groups, our analyses based on our cleanest samples with our most comprehensive specifications

(the switcher models in Columns 4 and 5 of Table 3 and high-risk vehicles in Columns 1 and 2 of Table 4) indicate that service and repair facilities exhibit more leniency than gas retailers, which have the least incentive to falsely pass customers.

In studying the impact of governance, we find that branded facilities and subsidiaries are less lenient than independent facilities, consistent with our hypotheses that the former governance structures increase the cost of failing to enforce regulations. We observe these same effects of governance and scope on leniency when we limit our analysis to changes in test results within particular vehicles as they change facilities. The panel nature of these within-vehicle analyses resolves much of the omitted-variable bias and most of the endogenous selection problems inherent in cross-sectional analyses.

One interesting result that deserves further investigation is the potential impact of AAA certification on leniency. Despite the small number of AAA-certified facilities in our sample (44), which limits the statistical power of our empirical analysis, we find evidence that AAA certification is associated with greater monitoring stringency in our switcher and high-risk samples. A growing literature argues that third-party certification of operational process conformance can reduce socially harmful activities (Potoski and Prakash 2005b; Levine and Toffel 2010; Short and Toffel 2010). Unfortunately, data limitations—specifically our small number of AAA facilities (44) and our inability to obtain from AAA the precise certification dates of the facilities in our sample—prevent us from distinguishing selection effects from treatment effects. We encourage future research to tease apart the extent to which third-party certification attracts and identifies facilities with more stringent monitoring practices (a selection effect) and/or encourages more stringent monitoring (a treatment effect). In addition, further research could explore the effectiveness of corporations requiring their franchises and subsidiaries to pursue third-party certification (Darnall 2006) to ensure that their private monitoring stringency adheres to corporate standards.

Contributions

Our research contributes to the literatures exploring the performance implications of private governance, quality management, and industry self-regulation. Whereas a large literature examines franchising as a governance form (LaFontaine 1992; Chung and Kalnins 2001; Kalnins and Mayer 2004; Mitsuhashi, Shane, and Sine 2008), the performance implications of franchising have received much less attention (Barthélemy 2008). A few studies have examined the performance of franchises and chains in terms of firm survival (Shane 1996, 1998; Shane and Foo 1999), sales growth (Sorenson and Sorensen 2001; Yin and Zajac 2004) and returns on assets and sales (Barthélemy 2008). In one of the few studies besides ours that focuses on the operational performance implications of subsidiaries and franchising brands, Jin and Leslie (2009) found that restaurants of these types exhibited better hygiene than independent restaurants did, suggesting that, as in our results, the former exhibited more stringent operational control. Others found that subsidiaries exhibited better compliance with labor laws (Ji and Weil 2009) and greater production-performance variability (Hsieh et al. 2010). These studies, in combination with our own, suggest that both incentives and operational control matter. But they also suggest that future research is needed to theoretically and empirically distinguish how these factors affect and are affected by firms' selection of governance structures and their performance implications. Since governance affects a firm's incentives and its capabilities—both of which affect its behavior—separating these two is an important step towards fully explaining the mechanisms through which governance influences behavior.

Our results also contribute to the broad literature on corporate governance in financial auditing (e.g., Larcker and Richardson 2004). Major scandals such as Enron and WorldCom led to the Sarbanes-Oxley Act's (2002) substantial policy changes in how auditors are appointed and their scope of activities. Research in accounting has found that governance is highly related to oversight quality and fraud reduction (Farber 2005; Abbott et al. 2004), but recent work suggests that some problems may still exist. For example, changes dictated by Sarbanes-Oxley that require auditors to be chosen and managed by boards of directors, rather than by top management, were designed to reduce the likelihood of side

payments for leniency, but recent interviews suggest some continued influence by management (Cohen et al. 2010). We believe our setting, which at face value seems radically different from auditing but possesses some similar characteristics, can help inform this large and influential literature.

Our findings also contribute to the literature on industry self-regulation, in which studies have found superior operational performance among facilities that had been independently certified to international process management standards (Dasgupta, Hettige, and Wheeler 2000; King and Lenox 2001; Levine and Toffel 2010; Potoski and Prakash 2005a, 2005b; Toffel 2006). Our findings build on these studies by revealing that buyers and regulators can also rely on other private governance mechanisms, including subsidiary and franchise relationships, as credible indicators that facilities are engaging in more stringent private regulatory monitoring. We also found limited evidence that AAA-certified facilities engage in particularly stringent monitoring (Columns 4 and 5 of Table 3), despite the fact that, in contrast to most independently certified standards, AAA both develops the standards behind its AAA Approved Auto Repair Network and exclusively conducts the site visits to verify compliance. Our results reveal that, at least in this context, the same organization can both promulgate a process standard and conduct certifications in a manner that, despite concerns about conflicts of interest, can result in certified establishments outperforming non-certified establishments.

Limitations and Future Research

We acknowledge several limitations to our study. First, we note that our real-world empirical setting, like most, limits our ability to test the full range of our theoretical predictions. We cannot, for example, directly observe levels of customer fear of moral hazard in cross-selling and must therefore infer it based on past studies of the automotive industry. Similarly, we can only observe broad classifications of governance and firm scope. While industry evidence and the literature provide clear predictions about how the governance and scope designations in our models fit our theory, other industries might provide more substantial variation along these dimensions than we are able to observe here.

Although our empirical models control for many characteristics of facilities, vehicles, testing conditions, and—in most cases—unobservable time-invariant vehicle factors, we cannot observe all the factors that influence vehicle owners’ decisions about which facilities to patronize. If unobserved factors that affect this decision are also correlated with the propensity to receive lenient monitoring, our results would be vulnerable to omitted-variable bias. However, if such factors exist in our context, we believe the most likely scenario would result in a bias against, not in favor of, our hypothesized effects. Specifically, we suspect that vehicle owners seeking leniency would be especially likely to patronize independent facilities in order to avoid the private governance oversight that characterizes (a) subsidiaries and branded franchisees and (b) service and repair facilities and car dealers. Indeed, supplemental analysis (described in the Appendix) indicates that owners of failing vehicles were especially likely to seek subsequent inspections at independently governed service and repair facilities and independently governed car dealers. To the extent that owners seeking leniency (presumably with more poorly maintained vehicles) are fleeing to independents, subsidiary and branded facilities would be seeing vehicles in better-than-average condition and would therefore be especially *unlikely to fail* vehicles. This scenario would represent a bias against our empirical results, which instead found that subsidiaries and branded facilities were in fact especially *unlikely to pass* vehicles.

We also note that we cannot observe or control for the endogenous decisions by firm owners to choose their governance mode and organizational scope. The identification concern is that more ethical or law-abiding owners would choose to brand themselves or to select into markets where we observe lower leniency, such as gas retailing. While we have no reason to suspect this is driving our results, we are unable to rule out this alternative explanation. Only 2.5% of our facilities change governance mode, which does not allow us to model this selection process.

We must also note that, in our setting, all firms and locations are regulated by a single agency, which intensifies the negative reputational spillovers of excessive leniency detected at any one location. It is unclear how strong such spillovers would be if subsidiaries or locations of the same brand were regulated by different agencies. In our setting, this could involve emissions testing gas retailers that share

the same brand but operate in different American states (since inspections are regulated at the state rather than at the federal level). We suspect that these spillovers would be weaker than those within a particular state agency's regulatory span, much as we believe that positive spillovers through customers' referrals are weak, although this is certainly open to future investigation. Furthermore, emissions testing fraud in New York is prosecuted by the office of the New York State Attorney General, which has shown great willingness to challenge firms that violate criminal and environmental law.²² This suggests that large brands or corporate parents (and their substantial legal resources) would not deter prosecution in our setting, though they might well do so in other settings.

Although our research has focused on leniency, we have noted that some firms might have incentives to be overly stringent and fraudulently fail vehicles in order to profit from performing unnecessary repairs. Although we acknowledge the possibility that our data might include some overly stringent emissions tests, the market's competitive dynamics and the fact that investigations consistently find fraud in the form of leniency suggest that over-stringency is uncommon. Furthermore, the capitation on emissions repairs in our focal state limits the incentive for such over-stringency while competition in the testing market enables owners of such inappropriately failed vehicles to find other facilities that would (correctly) pass them. Facilities could only profit from fraudulently failing vehicles owned by customers unaware of—or unwilling to make use of—other testing facilities. Even owners whose vehicles have been falsely failed are only likely to repair their cars if the cost is low, since they have other alternatives, including selling their vehicles in regions with less stringent emissions standards or no standards at all. As a result, nearly all fraud exposés in the vehicle emissions testing market concern monitoring leniency rather than over-stringency (e.g., Lambert 2000; Groark 2002; States News Service 2009a, 2009b; Navarro 2010; Roosevelt 2010; U.S. Department of Justice 2010, 2011). For example, of the covert tests New York conducted in 2003 using vehicles rigged to fail, 40% (117 of 293) resulted in false passes

²² Eliot Spitzer, who was New York Attorney General during our time period, was famous for his aggressiveness in pursuing corporate offenders. The EPA, which is charged with enforcing these regulations if states do not, also famously prosecuted Ford Motor Company for emissions testing fraud in the 1970s (Fisse and Braithwaite 1983).

(New York State Department of Environmental Conservation, 2004b). After officials in Salt Lake County, Utah conducted 4,352 covert and overt investigations of 320 car dealerships and service facilities that performed emission tests, “the major violations usually involved people testing one car in the place of another” to pass vehicles that ought to have failed (Groark 2002: 4).

Future research could examine markets that risk overly stringent private monitoring and explore whether private governance and organizational scope have similar or different impacts on lenient versus overly stringent monitoring. One potential avenue for this research is to follow the literature on franchising in examining those firms where repeat business is unlikely (e.g., Jin and Leslie 2009). Unfortunately, this is difficult to do in our setting, since we have little information on car ownership and minimal geographic variation. Similarly, it is important to note that, in many markets with privatized monitoring, monitors will likely err on the side of stringency to protect themselves from devastating reputation consequences. Markets in which the customers themselves value stringency—such as the inspection of elevators, brakes, and boilers—are also unlikely to suffer from fraudulent leniency.

Finally, while we described several reasons why, in our empirical context, an establishment’s reputation for leniency was unlikely to stimulate demand at its corporate sibling due to the illicit nature of leniency and the geographic constraints of most customers (who tend to seek local testing facilities), such positive spillover benefits might arise in other contexts. Co-owned locations could potentially benefit as customers shared their experience of leniency with others in their social or professional networks, especially when a customer receiving leniency at one location can easily share this with geographically distant social contacts (e.g., through online social networking), who might then seek leniency at the monitoring firm’s other locations. In our setting, this constitutes encouraging distant contacts to solicit fraudulent behavior and also requires understanding each location’s ownership. Further research is warranted to investigate circumstances in which subsidiary customers are prone to sharing such information, as this could counteract the subsidiary owner’s tendency to deter leniency.

Conclusion

As firms and government organizations increasingly outsource manufacturing and services, managers increasingly need to decide how to assure the quality of these products and services. Managers face an expanding number of options, from relying on contracts and the supplier's brand to employing a team of quality specialists who visit supplier factories to hiring third parties to perform factory audits.

Our study reveals systematic leniency among those third-party monitors whose scope of services results in conflicts of interest because there are profitable opportunities in extending leniency to engender customer loyalty. We also found leniency among third-party monitors that lacked governance mechanisms that enhance internal controls. Specifically, we showed that the quality of monitoring is higher at subsidiaries and branded affiliates than at independent facilities.

By pointing managers to the hazards of third-party monitoring, our findings can guide their efforts to efficiently scrutinize monitors to enhance their effectiveness. Governments can apply these insights well beyond the environmental context we examined, to a wide array of domains in which outsourcing is considered, such as hiring contractors to assess the integrity of bridges and tunnels, food safety, financial accounting accuracy, and conformity to public school standards. By exposing particular types of organizational scope and governance that are especially prone to leniency, our results can also help managers and policy makers prioritize their investigations of third-party monitors.

REFERENCES

- Abbott, L., S. Parker, G. Peters. 2004. Audit committee characteristics and restatements. *Auditing: A Journal of Practice & Theory* **23**(1) 69-87.
- American Automobile Association. 2010. The AAR approval process. <http://partner.aaa.biz/portal/site/auto/menuitem.d4bee97712854315dd1b2110089fd3f8/>.
- Antle, R. 1984. Auditor independence. *Journal of Accounting Research* **22**(1) 1-20.
- Banker, R. D., I. Khosla, K. K. Sinha. 1998. Quality and competition. *Management Science* **44**(9) 1179-1192.
- Barthélemy, J. 2008. Opportunism, knowledge, and the performance of franchise chains. *Strategic Management Journal* **29**(13) 1451-1463.
- Baum, J. A. C., A. M. McGahan. 2009. Outsourcing war: The evolution of the private military industry after the Cold War. Working paper, University of Toronto Rotman School of Management, Toronto.
- Becker, B., T. Milbourn. 2011. How did increased competition affect credit ratings? *Journal of Financial Economics* **101**(3) 493-514.
- Bénabou, R., J. Tirole. 2006. Incentives and prosocial behavior. *American Economic Review* **96**(5) 1652-1678.
- Bennett, V., L. Pierce, J. Snyder, M. Toffel. 2012. Competition and illicit quality. Working paper, Harvard Business School, Boston.
- Bertrand, M., S. Djankov, R. Hanna, S. Mullainathan. 2007. Obtaining a driver's license in India: An experimental approach to studying corruption. *Quarterly Journal of Economics* **122**(4) 1639-1676.
- Bolton, P., X. Freixas, J. Shapiro. 2012. The credit ratings game. *Journal of Finance* **67**(1) 85-112.
- Boyd, C. 2004. The structural origins of conflicts of interest in the accounting profession. *Business Ethics Quarterly* **14**(3) 377-398.
- Bradach, J. L. 1997. Using the plural form in the management of restaurant chains. *Administrative Science Quarterly* **42**(2) 276-303.
- Brams, J. B. 1999. Franchisor liability: Drafting around the problems with franchisor control. *Oklahoma City University Law Review* **24** 65-79.
- Cabral, S., P. Furquim, S. Lazzarini. 2010. Private operation with public supervision: Evidence from hybrid modes of governance in prisons. *Public Choice* **145**(1-2) 281-293.
- Cabral, S., P. Furquim, S. Lazzarini. 2012. Private entrepreneurs in public services: A longitudinal examination of outsourcing and statization of prisons. *Strategic Entrepreneurship Journal*. Forthcoming.
- California Bureau of Automotive Repair. 2011. Final statement of reasons for Article 5.5 of Chapter 1 of Division 33 of Title 16 of the California Code of Regulations. June 10. http://www.smogcheck.ca.gov/80_BARResources/05_Legislative/RegulatoryActions/STAR%20Program%20FSOR_10_28_2011%20Final.pdf
- Chung, W., A. Kalnins. 2001. Agglomeration effects and performance: A test of the Texas lodging industry. *Strategic Management Journal* **22**(10) 969-997.
- Cohen, J., G. Krishnamoorthy, A. Wright. 2010. Corporate governance in the post-Sarbanes-Oxley era: Auditors' experiences. *Contemporary Accounting Research* **27**(3) 751-786.
- Cohen, M. A. 1987. Optimal enforcement strategy to prevent oil spills: An application of a principal-agent model with moral hazard. *Journal of Law and Economics* **30** 23-51.
- Cohen, M. A. 2000. Empirical research on the deterrence effect of environmental monitoring and enforcement. *Environmental Law Reporter* **30** 10245-10252.
- Cohn, D. 1992. EPA announces plan to expand, toughen auto emissions testing. *Washington Post* (July 14) A1.
- Corbett, C. J., M. J. Montes-Sancho, D. A. Kirsch. 2005. The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Science* **51**(7) 1046-1059.
- Corona, C., R. S. Randhawa. 2010. The auditor's slippery slope: An analysis of reputational incentives. *Management Science* **56**(6) 924-937.
- Coupé, T. 2005. Bias in conditional and unconditional fixed effects logit estimation: A correction. *Political Analysis* **13**(3) 292-295.

- Crosby, P. 2011. Emissions inspectors plead guilty to violating Clean Air Act. United States Attorney's Office. May 4. <http://www.justice.gov/usao/gan/press/2011/05-04-11.html>.
- Cutler, D. M., R. S. Huckman, J. T. Kolstad. 2010. Input constraints and the efficiency of entry: Lessons from cardiac surgery. *American Economic Journal: Economic Policy* **2**(1) 51-76.
- Dal Bó, E. 2006. Regulatory capture: A review. *Oxford Review of Economic Policy* **22**(2) 203-225.
- Darby, M., E. Karni. 1973. Free competition and the optimal amount of fraud. *Journal of Law and Economics* **16**(1) 67-88.
- Darnall, N. 2006. Why firms mandate ISO 14001. *Business & Society* **45**(3) 354-381.
- Darnall, N., S. Sides. 2008. Assessing the performance of voluntary environmental programs: Does certification matter? *Policy Studies Journal* **36**(1) 95-117.
- Dasgupta, S., H. Hettige, D. Wheeler. 2000. What improves environmental compliance? Evidence from Mexican industry. *Journal of Environmental Economics and Management* **39**(1) 39-66.
- Davidson, W., III, P. Jiraporn, P. DaDalt. 2006. Causes and consequences of audit shopping: An analysis of auditor opinion, earnings management, and auditor changes. *Quarterly Journal of Business and Economics* **45**(1) 69-87.
- Drugov, M. 2010. Competition in bureaucracy and corruption. *Journal of Development Economics* **92**(3) 107-114.
- Dun & Bradstreet. 2012. Paydex Score: Business Definition. <http://smallbusiness.dnb.com/glossaries/paydex-score/11815414-1.html>.
- Emons, W. 1997. Credence goods and fraudulent experts. *RAND Journal of Economics* **28**(1) 107-119.
- Epple, D., M. Visscher. 1984. Environmental pollution: Modeling occurrence, detection and deterrence. *Journal of Law and Economics* **27** 29-60.
- Ernst, M., J. Corless, R. Greene-Roesel. 2003. *Clearing the Air: Public Health Threats from Cars and Heavy Duty Vehicles—Why We Need to Protect Federal Clean Air Laws*. Surface Transportation Policy Partnership, Washington, DC. <http://www.transact.org/report.asp?id=227>.
- Fan, S., C. Lin, D. Treisman. 2009. Political decentralization and corruption: Evidence from around the world. *Journal of Public Economics* **93**(1-2) 14-34.
- Farber, D. 2005. Restoring trust after fraud: Does corporate governance matter? *Accounting Review* **80**(2) 539-561.
- First Research. 2010. Automotive repair shops. First Research, Austin, TX. January 25, 2010. <http://www.firstresearch.com>.
- Fisse, B., J. Braithwaite. 1983. *The Impact of Publicity on Corporate Offenders*. State University of New York Press, Albany, NY.
- Franzen, E. U. 2008. Emissions testing: Promising idea, but? Before DOT officials can change the rules on how vehicles are tested in the state, the Legislature has to change the law. *Milwaukee Journal Sentinel* (July 26) 8.
- Freeman, J., M. Minow, eds. 2009. *Government by Contract: Outsourcing and American Democracy*. Harvard University Press, Cambridge, MA.
- Gino, F., L. Pierce. 2010. Robin Hood under the hood: Wealth-based discrimination in illicit customer help. *Organization Science* **21**(6) 1176-1194.
- Greene, W. 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal* **7**(1) 98-119.
- Groark, V. 2002. An overhaul for emissions testing. *New York Times* (June 9) 4.
- Harrington W., V. McConnell. 1999. Coase and car repair: Who should be responsible for emissions of vehicles in use? Discussion Paper No. 99-22, Resources for the Future, Washington, DC.
- Hart, O., A. Shleifer, R. Vishny. 1997. The proper scope of government: Theory and an application to prisons. *Quarterly Journal of Economics* **112**(4) 1127-1161.
- He, J., J. Qian, P. Strahan. 2011. Credit ratings and the evolution of the mortgage-backed securities market. *American Economic Review: Papers & Proceedings* **101**(3) 131-135.
- Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* **10**(1) 74-91.
- Holmstrom, B., P. Milgrom. 1991. Multi-task principal-agent problems: Incentive contracts, asset ownership and job design. *Journal of Law, Economics, & Organization* **7** 24-52.

- Hsieh, C., S. G. Lazzarini, J. A. Nickerson, M. C. Laurini. 2010. Does ownership affect the variability of the production process? Evidence from international courier services. *Organization Science* **21**(4) 892-912.
- Hubbard, T. N. 1998. An empirical examination of moral hazard in the vehicle inspection market. *RAND Journal of Economics* **29**(2) 406-426.
- Hubbard, T. N. 2002. How do consumers motivate experts? Reputational incentives in an auto repair market. *Journal of Law and Economics* **45**(2) 437-468.
- Huckman, R. S., G. P. Pisano. 2006. The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science* **52**(4) 473-488.
- Jamal, K. 2008. Building a better audit profession. *Accounting Perspectives* **7**(2) 123-126.
- Jensen, M. C., W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* **3**(4) 305-360.
- Jensen, R. 1992. Reputational spillovers, innovation, licensing, and entry. *International Journal of Industrial Organization* **10**(2) 193-212.
- Ji, M., D. Weil. 2009. Does ownership structure influence regulatory behavior? The impact of franchisee free-riding on labor standards compliance. Working paper, Boston University, Boston.
- Jin, G., P. Leslie. 2009. Reputational incentives for restaurant hygiene. *American Economic Journal: Microeconomics* **1**(1) 237-267.
- Kalnins, A., K. Mayer 2004. Franchising, ownership, and experience: A study of pizza restaurant survival. *Management Science* **50**(12) 1716-1728.
- Katz, E. 2001. Bias in conditional and unconditional fixed effects logit estimation. *Political Analysis* **9**(4) 379-384.
- Khalil, F., J. Lawaree. 2006. Incentives for corruptible auditors in the absence of commitment. *Journal of Industrial Economics* **54**(2) 269-291.
- King, A. A., M. J. Lenox. 2001. Lean and green? An empirical examination of the relationship between lean production and environmental performance. *Production Operations Management* **10**(3) 244-256.
- King, A., M. Lenox, A. Terlaak. 2005. The strategic use of decentralized institutions: Exploring certification with the ISO 14001 management standard. *Academy of Management Journal* **48**(6) 1091-1106.
- Klein, B., R. G. Crawford, A. A. Alchian. 1978. Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics* **21**(2) 297-326.
- Klein, B., K. Leffler. 1981. The role of market forces in assuring contractual performance. *Journal of Political Economy* **89**(4) 615-641.
- Kreps, D. 1990. Corporate culture and economic theory. J. Alt, K. Shepsle, eds. *Perspectives on Positive Political Economy*. Cambridge University Press, Cambridge, UK, 90-143.
- Laffont, J.-J., T. N'Guessan. 1999. Competition and corruption in an agency relationship. *Journal of Development Economics* **60**(2) 271-295.
- Laffont, J.-J., J. Tirole. 1993. Cartelization by regulation. *Journal of Regulatory Economics* **5**(2) 111-130.
- LaFontaine, F. 1992. Agency theory and franchising: Some empirical results. *RAND Journal of Economics* **23**(2) 263-283.
- Lafontaine, F., R. D. Blair. 2009. The evolution of franchising and franchise contracts: Evidence from the United States. *Entrepreneurial Business Law Journal* **3**(2) 381-434.
- LaFontaine, F., K. L. Shaw. 2005. Targeting managerial control: Evidence from franchising. *RAND Journal of Economics* **36**(1) 131-150.
- Lambert, C. 2000. Payoffs alleged as emissions tests end. *Palm Beach Post* (June 30) 1A.
- Lamure, M., S. Hoffmann, and L. Flees. 2009. Auto survey: Improving after-sales service helps retain customers. *Shanghai Daily* (February 4).
http://www.bain.com/bainweb/publications/publications_detail.asp?id=26889&menu_url=publication_s_results.asp.
- Laplante, B., P. Rilstone. 1996. Environmental inspections and emissions in the pulp and paper industry in Quebec. *Journal of Environmental Economics and Management* **31**(1) 19-36.
- Larcker, D. F., S. A. Richardson. 2004. Fees paid to audit firms, accrual choices, and corporate governance. *Journal of Accounting Research* **42**(3) 625-658.

- Lazare, D. 1980. Jersey weighs its shift to private inspections. *New York Times* (October 26) S21.
- Leamer, E. E. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. Wiley, New York.
- Lennox, C., J. Pittman. 2010. Auditing the auditors: Evidence on the recent reforms to the external monitoring of audit firms. *Journal of Accounting and Economics* **49**(1-2) 84-103.
- Levin, J., S. Tadelis. 2010. Contracting for government services: Theory and evidence from U.S. cities. *Journal of Industrial Economics* **58**(3) 507-541.
- Levine, D. I., M. W. Toffel. 2010. Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. *Management Science* **56**(6) 978-996.
- Levitt, A. 2000. Renewing the covenant with investors. Speech by Chairman Arthur Levitt, U.S. Securities and Exchange Commission, at New York University's Center for Law and Business. U.S. Securities and Exchange Commission, New York.
http://www.financialexecutives.org/download/SEC_Speech_5-10-00.pdf.
- Luca, M. 2011. Reviews, reputation, and revenue: The case of Yelp.com. Working paper, Harvard Business School, Boston.
- Mahoney, J. T., A. M. McGahan, C. N. Pitelis. 2009. The interdependence of private and public interests. *Organization Science* **20**(6) 1034-1052.
- Mayer, K. J. 2006. Spillovers and governance: An analysis of knowledge and reputational spillovers in information technology. *Academy of Management Journal* **49**(1) 69-84.
- Mayer, K. J., J. A. Nickerson, H. Owan. 2004. Are supply and plant inspections complements or substitutes? A strategic and operational assessment of inspection practices in biotechnology. *Management Science* **50**(8) 1064-1081.
- Minow, M. 2005. Outsourcing power: How privatizing military efforts challenges accountability, professionalism, and democracy. *Boston College Law Review* **46**(5) 989-1026.
- Mitsubishi, H., S. Shane, W. D. Sine. 2008. Organization governance form in franchising: Efficient contracting or organizational momentum? *Strategic Management Journal* **29**(10) 1127-1136.
- Moore, D. A., P. E. Tetlock, L. Tanlu, M. H. Bazerman. 2006. Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* **31**(1) 10-29.
- National Association of Convenience Stores. 2009. How do retailers get—and sell—gasoline? February 2.
http://www.nacsonline.com/NACS/Resources/campaigns/GasPrices_2009/Pages/HowRetailersGetSellGas.aspx.
- National Research Council, Committee on Vehicle Emission Inspection and Maintenance Programs. 2001. *Evaluating Vehicle Emissions Inspection and Maintenance Programs*. National Academy Press, Washington, DC.
- Navarro, M. 2010. Vehicle tests on emissions were faked. *New York Times* (February 19) A-21.
- Nelson, P. 1970. Information and consumer behavior. *Journal of Political Economy* **78**(2) 311-329.
- New Strategist. 2007. *Household Spending: Who Spends How Much on What*. New Strategist Publications, Ithaca, NY.
- New York State Department of Environmental Conservation. 2004a. *2003 I/M Program Annual Report: Conclusion*. <http://www.dec.ny.gov/chemical/23734.html>.
- New York State Department of Environmental Conservation. 2004b. *2003 I/M Program Annual Report: Data Analysis*. <http://www.dec.ny.gov/chemical/23735.html>.
- New York State Department of Environmental Conservation. 2004c. *2003 I/M Program Annual Report: Introduction*. <http://www.dec.ny.gov/chemical/23736.html>.
- New York State Department of Environmental Conservation. 2010. 40 facilities cited for falsifying 20,000+ vehicle emissions inspections. February 18. <http://www.dec.ny.gov/press/62821.html>.
- New York State Department of Environmental Conservation. 2011. *2010 Enhanced I/M Program Annual Report*. <http://www.dec.ny.gov/chemical/76285.html>.
- New York State Department of Motor Vehicles. 2004. *Motor Vehicle Inspection Regulations as of November 2004*, section 79.17. http://www.dec.ny.gov/docs/air_pdf/vipsipapp2.pdf.

- New York State Department of Motor Vehicles. 2011. *Motor Vehicle Inspection Regulations as of July 2011*. <http://www.nydmv.state.ny.us/forms/cr79.pdf>.
- Nickerson, J. A., B. S. Silverman. 2003. Why aren't all truck drivers owner-operators? Asset ownership and the employment relation in interstate for-hire trucking. *Journal of Economics and Management Strategy* **12**(1) 91-118.
- Oliva, P. 2012. Environmental regulations and corruption: Automobile emissions in Mexico City. Working paper, University of California at Santa Barbara, Santa Barbara, CA.
- Olmstead, A. L., P. Rhode. 1985. Rationing without government: The West Coast gas famine of 1920. *American Economic Review* **75**(5) 1044-1055.
- Perkins, H. C., S. J. Yatchak, G. M. Hadfield. 2010. Franchisor liability for acts of the franchisee. *Franchise Law Journal* **29**(3) 174-207.
- Pesendorfer, W., A. Wolinsky. 2003. Second opinions and price competition: Inefficiency in the market for expert advice. *Review of Economic Studies* **70**(2) 417-437.
- Pierce, L., J. Snyder. 2008. Ethical spillovers in firms: Evidence from vehicle emissions testing. *Management Science* **54**(11) 1891-1903.
- Pierce, L., J. Snyder. 2012. Discretion and manipulation by experts: Evidence from a vehicle emissions policy change. *B.E. Journal of Economic Analysis and Policy*. Forthcoming.
- Potoski, M., A. Prakash. 2005a. Covenants with weak swords: ISO 14001 and facilities' environmental performance. *Journal of Policy Analysis and Management* **24**(4) 745-769.
- Potoski, M., A. Prakash. 2005b. Green clubs and voluntary governance: ISO 14001 and firms' regulatory compliance. *American Journal of Political Science* **49**(2) 235-248.
- Reichelstein, S. 1992. Constructing incentive schemes for government contracts: An application of agency theory. *Accounting Review* **67**(4) 712-731.
- Rey, P., B. Salanie. 1990. Long-term, short-term and renegotiation: On the value of commitment in contracting. *Econometrica* **58**(3) 597-619.
- Richman, B. 2004. Firms, courts, and reputation mechanisms: Towards a positive theory of private ordering. *Columbia Law Review* **104**(8) 2328-2368.
- Roosevelt, M. 2010. "Widespread fraud" in California's smog test program. *Los Angeles Times* (February 25). <http://articles.latimes.com/2010/feb/25/local/la-me-smogtest25-2010feb25/2>.
- Rule, S. 1978. Carey proposes an auto-exhaust inspection plan. *New York Times* (June 18) 52.
- Sarbanes-Oxley Act. 2002. 107th Congress of the United States of America, Second Session. <http://www.govtrack.us/congress/bills/107/hr3763/text>.
- Schneider, H. 2011. Agency problems and reputation in expert services: Evidence from auto repair. *Journal of Industrial Economics*. Forthcoming.
- Sclar, E. 2000. *You Don't Always Get What You Pay For: The Economics of Privatization*. Cornell University Press, Ithaca, NY.
- Seamans, R. C. 2011. Fighting city hall: Entry deterrence and new technology deployment in local cable TV markets. *Management Science*. Forthcoming.
- Seifter, M. 2009. Rent-a-regulator: Design and innovation in environmental decision making. J. Freeman, M. Minow, eds. *Government by Contract: Outsourcing and American Democracy*. Harvard University Press, Cambridge, MA, 93-109.
- Shane, S. A. 1996. Hybrid organizational arrangements and their implications for firm growth and survival: A study of new franchisors. *Academy of Management Journal* **39**(1) 216-234.
- Shane, S. A. 1998. Making new franchise systems work. *Strategic Management Journal* **19**(7) 697-707.
- Shane, S. A., M. Foo. 1999. New firm survival: An institutional explanation for franchisor mortality. *Management Science* **45**(2) 142-159.
- Shimshack, J. P., M. B. Ward. 2005. Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management* **50** 519-540.
- Shleifer, A., R. W. Vishny. 1993. Corruption. *Quarterly Journal of Economics* **108**(3) 599-617.
- Short, J. L., M. W. Toffel. 2008. Coerced confessions: Self-policing in the shadow of the regulator. *Journal of Law, Economics and Organization* **24**(1) 45-71.
- Short, J. L., M. W. Toffel. 2010. Making self-regulation more than merely symbolic: The critical role of the legal environment. *Administrative Science Quarterly* **55**(3) 361-396.

- Sorenson, O., J. B. Sorensen. 2001. Finding the right mix: Franchising, organizational learning, and chain performance. *Strategic Management Journal* **22**(6-7) 713-724.
- Spence, M. 1977. Consumer misperceptions, product failure and product liability. *Review of Economic Studies* **44**(3) 561-572.
- States News Service. 2009a. Los Angeles smog check stations face criminal charges for illegally certifying cars. September 3.
- States News Service. 2009b. MassDEP environmental strike force investigation results in penalties, suspension of licenses at six auto emission inspection stations. October 16.
- Sultan, A. 2008. Lemons hypothesis reconsidered: An empirical analysis. *Economics Letters* **99**(3) 541-544.
- Taylor, C. R. 1995. The economics of breakdowns, checkups, and cures. *Journal of Political Economy* **103**(1) 53-74.
- Terlaak, A., A. King. 2006. The effect of certification with the ISO 9000 quality management standard: A signaling approach. *Journal of Economic Behavior and Organization* **60**(4) 579-602.
- Thornton, D., N. Gunningham, R. A. Kagan. 2005. General deterrence and corporate environmental behavior. *Law and Policy* **27** 262-288.
- Tirole, J. 1986. Hierarchies and bureaucracies: On the role of collusion in organizations. *Journal of Law, Economics, & Organization* **2**(2) 181-214.
- Tirole, J. 1992. Collusion and the theory of organizations. Laffont, J. J., ed. *Advances in Economic Theory, Sixth World Congress*. Vol. 2. Cambridge University Press, Cambridge, UK, 151-206.
- Toffel, M. W. 2006. Resolving information asymmetries in markets: The role of certified management programs. Harvard Business School Working Paper 07-023, Harvard Business School, Boston.
- U.S. Department of Justice. 2010. Press release: 10 Las Vegas men indicted for falsifying vehicle emissions tests. January 8. <http://www.justice.gov/opa/pr/2010/January/10-enrd-015.html>.
- U.S. Department of Justice. 2011. Last of 10 Las Vegas defendants sentenced for falsifying emissions test records, September 21. <http://www.epa.gov/compliance/resources/cases/criminal/highlights/2011/mccown-wm-09-21-11.pdf>.
- U.S. Environmental Protection Agency. 1994. *Clean Cars for Clean Air: Inspection and Maintenance Programs*. Fact Sheet OMS-14. EPA 400-F-92-016. Office of Mobile Sources. <http://www.epa.gov/oms/consumer/14-insp.pdf>.
- U.S. Environmental Protection Agency. 2007. Mobile source emissions: Past, present, and future. Office of Transportation and Air Quality. <http://www.epa.gov/otaq/invntory/overview/index.htm>.
- Voas, S., P. Shelly. 1995. Ridge kills new emissions tests. *Pittsburgh Post-Gazette* (October 19) A1.
- Wernerfelt, B. 1988. Umbrella branding as a signal of new product quality: An example of signaling by posting a bond. *RAND Journal of Economics* **19**(3) 458-466.
- Williamson, O. E. 1983. Credible commitments: Using hostages to support exchange. *American Economic Review* **73**(4) 519-540.
- Williamson, O. E. 1985. *The Economic Institutions of Capitalism*. Free Press, New York.
- Williamson, O. E. 1999. Public and private bureaucracies: A transaction cost economics perspective. *Journal of Law, Economics, & Organization* **15**(1) 306-342.
- Wulf, J. 2002. Internal capital markets and firm-level compensation incentives for division managers. *Journal of Labor Economics* **20**(2) S219-S262.
- Yin, X., E. J. Zajac. 2004. The strategy/governance structure fit relationship: Theory and evidence in franchising arrangements. *Strategic Management Journal* **25**(4) 365-383.

Figure 1: Predicted Impact of Scope on Emissions Testing Leniency

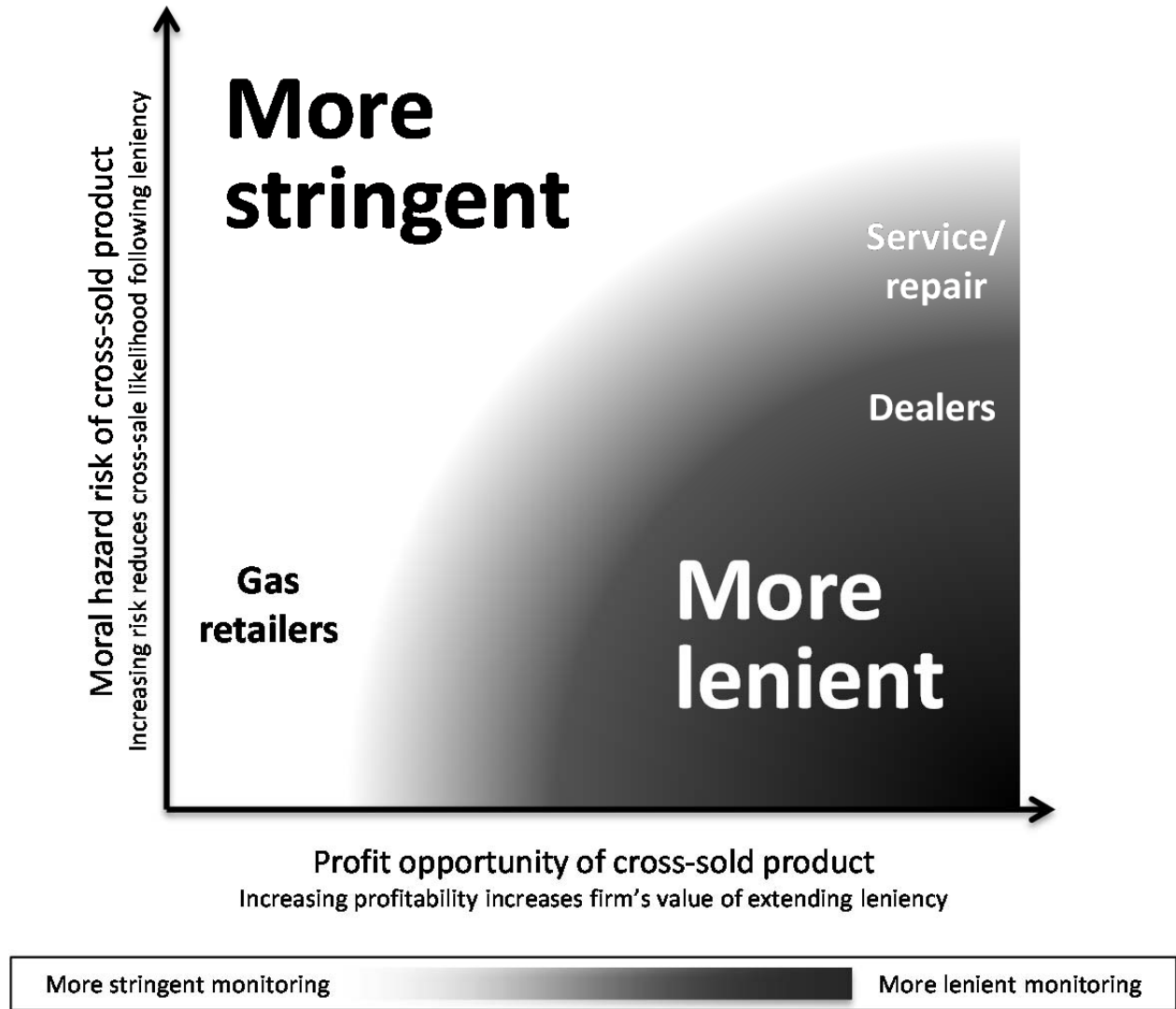


Table 1. Sample Description

Scope of facility activities	Organizational governance				Total
	Independent facilities	Branded and subsidiaries	Branded but not subsidiaries	Subsidiaries but not branded	
Gasoline retailers					
Unique facility-years	2,315	60	440	183	2,998
Inspections	1,186,157	217,20	235,054	78,248	1,521,179
Service/repair stations					
Unique facility-years	8,083	406	597	594	9,680
Inspections	3,841,260	195,592	279,772	232,750	4,549,374
Car dealers					
Unique facility-years	374	247	894	108	1,623
Inspections	126,178	71,501	230,747	32,297	460,723
Total					
Unique facility-years	10,772	713	1,931	885	14,301
Inspections	5,153,595	288,813	745,573	343,295	6,531,276

Table 2. Descriptive Statistics

Panel A. Summary Statistics

	Mean	SD	Min	Max
Passed emissions test	0.92	0.28	0	1
Gasoline retailer	0.23	0.42	0	1
Service and repair facility	0.70	0.46	0	1
Car dealer	0.07	0.26	0	1
Independent facility	0.79	0.41	0	1
Subsidiary facility	0.10	0.30	0	1
Subsidiary siblings	0.41	2.03	0	18
Subsidiary siblings (log)	0.13	0.45	0	2.94
Branded facility	0.16	0.37	0	1
Branded siblings	2.98	8.91	0	55
Branded siblings (log)	0.42	1.04	0	4.03
AAA-certified facility	0.02	0.12	0	1
Facility inspection volume (level)	86.17	55.63	1	454
Facility inspection volume (log plus 1)	4.28	0.63	0.69	6.12
Facility competition (level)	13.84	10.56	0	58
Facility competition (log)	2.43	0.79	0	4.08
Facility neighborhood's median household income	53,674	19,581	24,999	185,345
Facility neighborhood's median household income (log)	10.83	0.32	10.13	12.13
Vehicle odometer (10,000 miles)	9.59	5.60	0	100.00

Note: N=6,531,276 emissions tests

Panel B. Vehicle Statistics by Facility Type

Tests conducted at:	Gasoline retailers	Service/repair stations	Car dealers	Independent facilities	Subsidiaries	Branded facilities
Passed emissions test	0.92 {0.28}	0.91 {0.28}	0.95 {0.22}	0.91 {0.28}	0.91 {0.28}	0.92 {0.27}
Vehicle odometer (10,000 miles)	9.17 {5.23}	9.85 {5.75}	8.44 {4.97}	9.74 {5.67}	9.20 {5.44}	8.87 {5.11}
Vehicle age (years)	9.89 {3.44}	10.09 {3.45}	8.65 {3.26}	10.11 {3.46}	9.45 {3.38}	9.19 {3.36}
Vehicle weight (1,000 pounds)	3.13 {0.82}	3.13 {0.85}	3.18 {0.84}	3.14 {0.84}	3.12 {0.88}	3.12 {0.84}
Emissions tests (N)	1,521,179	4,549,374	460,723	5,153,595	632,108	1,034,386

Note: Figures reported are mean values, with standard deviation in curly brackets.

Panel C. Pairwise Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) Passed emissions test	1.00												
(2) Gasoline retailer	0.00	1.00											
(3) Service and repair facility	-0.02	-0.83	1.00										
(4) Car dealer	0.03	-0.15	-0.42	1.00									
(5) Independent facility	-0.01	-0.01	0.21	-0.35	1.00								
(6) Subsidiary facility	0.00	-0.06	-0.01	0.12	-0.63	1.00							
(7) Subsidiary siblings (log)	-0.01	-0.08	0.04	0.07	-0.55	0.87	1.00						
(8) Branded facility	0.01	0.02	-0.22	0.38	-0.84	0.27	0.34	1.00					
(9) Branded siblings (log)	0.01	0.05	-0.26	0.38	-0.79	0.24	0.32	0.94	1.00				
(10) AAA-certified facility	0.00	-0.07	0.08	-0.03	0.00	-0.02	-0.02	0.01	-0.01	1.00			
(11) Facility inspection volume (log plus 1)	0.03	0.03	-0.04	0.02	-0.01	0.01	0.04	0.03	0.03	-0.01	1.00		
(12) Facility competition (log)	0.00	-0.10	0.06	0.06	-0.04	0.06	0.08	0.02	0.02	-0.05	-0.01	1.00	
(13) Facility neighborhood's median household income (log)	0.02	0.08	-0.11	0.06	-0.13	0.05	0.06	0.15	0.14	0.00	-0.09	-0.03	1.00
(14) Vehicle odometer (10,000 miles)	-0.09	-0.04	0.07	-0.06	0.05	-0.02	-0.02	-0.06	-0.05	-0.02	0.04	0.01	-0.09

Note: N=6,531,276 emissions tests

Table 3. Impact of Scope and Governance on Leniency

Dependent variable: Passed emissions test

	(1)	(2)	(3)	(4)	(5)
Functional form	Logit	Logit	Logit	Conditional logit	Conditional logit
Standard errors clustered by	Facility	Facility	Facility	Vehicle	Vehicle
	Logit coefficients	Logit coefficients	Logit coefficients	Average marginal effect	Logit coefficients
Service and repair facility	0.036 [0.028]		0.038 [0.028]	0.003	0.113** [0.007]
Car dealer	0.234** [0.049]		0.282** [0.052]	0.020	0.209** [0.024]
Branded siblings (log)		-0.015 [0.010]	-0.030** [0.011]	-0.002	-0.049** [0.003]
Subsidiary siblings (log)		-0.127** [0.021]	-0.123** [0.021]	-0.009	-0.170** [0.006]
AAA-certified facility	-0.013 [0.067]	-0.020 [0.067]	-0.018 [0.067]	-0.001	-0.149** [0.024]
Facility competition (log)	-0.003 [0.017]	0.007 [0.016]	0.001 [0.016]	0.000	0.031** [0.004]
Facility inspection volume (log)	0.222** [0.022]	0.226** [0.021]	0.227** [0.022]	0.016	0.293** [0.004]
Facility neighborhood's median household income (log)	0.163** [0.050]	0.164** [0.049]	0.164** [0.049]	0.012	-0.027 [0.019]
Vehicle model fixed effects	Included	Included	Included		Absorbed
Vehicle conditional fixed effects					Absorbed
Vehicle-inspection month conditional fixed effects					Included
Odometer (level, squared, and cubed)	Included	Included	Included		Included
3-digit ZIP code fixed effects	Included	Included	Included		Included
Model age fixed effects	Included	Included	Included		Included
Inspection month fixed effects	Included	Included	Included		Included
Observations	6,531,276	6,531,276	6,531,276		1,144,318
Vehicles	2,748,039	2,748,039	2,748,039		342,764
Vehicle models	3,593	3,593	3,593		3,558
Facilities	3,530	3,530	3,530		3,484

Brackets contain robust standard errors clustered by facility for Models 1-3 and clustered by vehicle for Models 4-5 (for which clustering by facility was infeasible). ** p<0.01, * p<0.05, + p<0.10. The omitted firm-scope category is *gasoline retailer* and the omitted governance category is *independent*. Models 1-3 include fixed effects for the vehicle model, identified by the first eight digits of the Vehicle Identification Number (VIN). To facilitate model convergence and to ensure that the fixed effects do not introduce bias to unconditional logit estimates, Models 1-3 are estimated on a sample limited to vehicle models with at least 100 inspections and at least five emissions tests that failed. Model 4 includes conditional fixed effects for each vehicle, identified by VIN. Model 5 includes conditional fixed effects for each unique combination of vehicle (VIN) and inspection month, which proxies for vehicle-owner relationships since a particular vehicle is typically inspected in the same month every year except when the vehicle is sold, in which case the vehicle begins a new annual cycle of being testing in the month it was sold.

Table 4. Alternative Samples and Specifications

Dependent variable: Passed emissions test

	(1)	(2)	(3)
Functional form	Logit	Logit	Logit
Sample restriction	Note 1	Note 2	None
Service and repair facility	0.071+	0.108*	0.033
	[0.039]	[0.046]	[0.030]
Car dealer	0.226**	0.216*	0.182*
	[0.074]	[0.086]	[0.086]
Branded siblings (log)	-0.037*	-0.036*	-0.041*
	[0.015]	[0.017]	[0.017]
Subsidiary siblings (log)	-0.162**	-0.190**	-0.115**
	[0.026]	[0.032]	[0.034]
Branded siblings (log) × Test at service/repair station			0.006
			[0.024]
Branded siblings (log) × Test at car dealer			0.071*
			[0.034]
Branded siblings (log) × Subsidiary siblings (log)			-0.002
			[0.016]
AAA-certified facility	-0.181*	-0.193*	-0.016
	[0.076]	[0.095]	[0.068]
Facility competition (log)	0.019	0.025	-0.001
	[0.021]	[0.025]	[0.016]
Facility inspection volume (log)	0.238**	0.255**	0.227**
	[0.029]	[0.035]	[0.021]
Facility neighborhood's median household income (log)	0.045	0.022	0.163**
	[0.068]	[0.084]	[0.049]
Vehicle model fixed effects	Included	Included	Included
Odometer (level, squared, and cubed)	Included	Included	Included
3-digit ZIP code fixed effects	Included	Included	Included
Model age fixed effects	Included	Included	Included
Inspection month fixed effects	Included	Included	Included
Observations	369,102	153,490	6,531,276
Vehicles	227,489	138,088	2,748,039
Vehicle models	3,261	3,062	3,593
Facilities	3,309	3,090	3,530

Brackets contain robust standard errors clustered by facility. ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$. The omitted firm-scope category is *gasoline retailer* and the omitted governance category is *independent*. Like Models 1-3 in Table 3, all of the models in these tables include fixed effects for the vehicle model, identified by the first eight digits of the Vehicle Identification Number (VIN). To facilitate model convergence and to ensure that the fixed effects do not introduce bias to unconditional logit estimates, all of these models are estimated on a sample limited to vehicle models with at least 100 inspections and at least five emissions tests that failed.

Note 1. Sample limited to vehicles that failed an emissions test at least once in the prior three years and are thus at particular risk of failing the focal emissions test.

Note 2. We build this “at risk” sample by calculating the average annual deterioration (increase in emissions) of passing cars for each make/model group (8-digit VIN). Any car with results less than this amount above the failing threshold would be more than 50% likely to fail in the next year under normal deterioration rates.

APPENDIX: Failing Vehicles Erodes Customer Loyalty

We describe here our analysis that reveals that it can be costly for an inspection facility to fail a vehicle on its emissions test. Although the greatest financial benefits that inspection facilities can gain from leniency may be to engender customer loyalty that could enhance the cross-selling of cars, parts, repairs, and service, we are unable to observe such cross-selling activities. However, we can observe customer return rates for emissions tests, which we believe is a reasonable proxy for customer loyalty to the facility's other business activities (Hubbard 1998). Consequently, we estimate the following model to assess the impact of a vehicle failing an emissions test on the likelihood of that vehicle being brought back to the same facility for an emissions test the following year. The unit of analysis is the individual vehicle emissions inspection.

$$(A1) \text{ Stay}_{i,j,t} = F(\text{Fail}_{i,j,t-1}, \text{Scope}_{i,j,t-1}, \text{Governance}_{i,j,t-1}, \text{Certified}_{i,j,t-1}, \text{VehicleCtrls}_{j,t-1}, \text{TestCtrls}_{i,j,t-1}, \text{FacilityCtrls}_{i,j,t-1})$$

where $F(\cdot)$ is the logit function. $\text{Stay}_{i,j,t}$ is a dummy coded "1" if vehicle j at time t returned to the same inspection facility i at which it had been inspected the prior year and is coded "0" otherwise. $\text{Fail}_{i,j,t-1}$ is a dummy coded "1" when vehicle j failed its emission test at facility j the prior year ($t-1$) and coded "0" otherwise. $\text{Scope}_{i,j,t-1}$ represents a series of dummy variables that indicate whether the prior inspection facility i of vehicle j was a gasoline retailer, car dealer, or service and repair facility. $\text{Governance}_{i,j,t-1}$ represents a series of dummy variables that indicate whether the prior inspection facility i of vehicle j was a subsidiary, branded, or independent facility. $\text{Certified}_{i,j,t-1}$ is a dummy variable that indicates whether or not vehicle j 's prior inspection facility i was AAA-certified. $\text{VehicleCtrls}_{j,t-1}$ includes vehicle-model fixed effects for vehicle j as well as the vehicle's odometer level at its prior inspection and the squared and cubed values of that level. $\text{TestCtrls}_{i,j,t-1}$ refers to a full set of dummies that denote the inspection year and another full set of dummies that indicate inspection month. $\text{FacilityCtrls}_{i,j,t-1}$ includes the prior inspection

facility's monthly inspection volume averaged over the two months prior to the focal inspection as well as a full set of dummies for the facility's three-digit ZIP code.

We used logistic regression to estimate the likelihood that vehicles returned to the same facility for their subsequent inspection, given the ownership and governance of the original test facility and, critically, whether the vehicle passed or failed its prior inspection there. To facilitate model convergence and to pursue an even more conservative approach than required to ensure negligible bias from an unconditional fixed effects logit model (Katz 2001; Coupé 2005; Greene 2004), we limit this sample to vehicle models with at least 100 inspections. Due to the regression specifications (described below), our sample is further limited to inspection observations for which we observed the vehicle's subsequent inspection. In this sample of 3,788,045 inspections of 1,805,205 vehicles (including 3,060 vehicle models) conducted by 3,412 inspection facilities, seven percent of the inspections resulted in failure (mean = 0.07, SD = 0.26). Vehicles were tested at the same facility at which their previous test had taken place ($Stay_{i,j,t} = 1$) in just over half the observations (mean = 0.526, SD = 0.50).

We pursue a conservative approach by reporting standard errors clustered by facility.²³ The results, presented in Column 1 of Table A-1, indicate that failing an inspection significantly decreased the likelihood that a vehicle will return to the same station ($p < 0.01$). The average marginal effect indicates that failing an inspection decreases the likelihood of returning to a facility by 9.8 percentage points, an 18.6% decrease from the sample average return rate of 52.6%. These results indicate that facilities that fail vehicles are consistently less likely to enjoy repeat business in emissions testing, which implies that failing vehicles is costly to facilities. Furthermore, the full cost likely exceeds the cost of lost emissions testing business. Although we cannot observe repeat business in the firms' other products and services,

²³ As a robustness test, we also clustered to account for relationships between facilities. Specifically, we clustered at the headquarters level for all subsidiaries, at the brand level for all branded facilities (except branded subsidiaries, which were clustered by headquarters), and at the facility level for all independent facilities. The results were virtually identical.

we assume that a decrease in emissions testing returns is correlated with decreased return business in other services.

INSERT TABLE A-1 HERE

As another extension, we examined whether vehicles that failed emissions tests were particularly likely to shift toward the facility types we hypothesize to be most lenient: independently governed service and repair stations and car dealers. To test this, we predict a dependent variable coded “1” for inspections that occur at independently governed facilities that were either service and repair stations or car dealers and coded “0” otherwise (mean = 0.611, SD = 0.49). We used logistic regression to estimate this model, predicting this dichotomous variable using a specification otherwise identical to the model reported in Column 1. The results, reported in Column 2, indicate that vehicles that failed their prior emissions test were significantly more likely to be taken to an independently governed service and repair facility or an independently governed car dealer for their next inspection ($p < 0.01$) than to some other type of facility, even after controlling for the type of facility at which their prior test had occurred. Compared to the sample average probability of a vehicle choosing either of these two facility types (61.1%), this effect suggests that failing a test increases the probability by 1.7 percentage points to a 62.8% probability. This suggests that owners of vehicles failing emissions tests were especially likely to shift to independently governed service and repair stations and car dealers, the facility types we hypothesize to be most lenient.

Table A-1. Regression Results of Consumer Choice Models

Dependent variable:	(1) Vehicle returns to same inspection facility		(2) Vehicle inspected at independent service/repair shop or car dealer	
	Logit coefficients	Average marginal effects	Logit coefficients	Average marginal effects
Vehicle's prior test failed	-0.407** [0.013]	-0.098	0.097** [0.012]	0.017
Facility inspection volume	0.126** [0.044]	0.031	0.008 [0.036]	0.001
Observations (inspections)	3,788,045		3,787,034	
Vehicles	1,805,205		1,804,898	
Vehicle models	3,060		3,060	
Facilities	3,412		3,411	

Brackets contain robust standard errors clustered by facility. All models are estimated using logistic regression and also include vehicle-model fixed effects, odometer (level, squared, and cubed), and a full set of dummies for three-digit ZIP code, model age, inspection month, inspection year, and the prior test facility's scope, governance, and membership in the AAA Approved Auto Repair Network. Both samples include inspections of vehicle models with at least 100 inspections.

** p<0.01, * p<0.05, + p<0.10.