Profitable Misconduct, Corporate Governance, and Law Enforcement

Anat R. Admati, Nathan Atkinson, and Paul Pfleiderer*

This Draft: August 26, 2025

Abstract

This paper analyzes interactions between corporate governance and law enforcement practices, focusing on cases where deterrence is weak and harmful misconduct is profitable. We show how managerial compensation contracts, including stock-based compensation and insurance that shields managers from liability, systematically undermine enforcement efforts aimed both at the corporation and at managers. We also show that common enforcement policies such as discounting penalties when corporations voluntarily self-report or implement compliance programs can actually lead to more social harm by increasing the profitability of misconduct. Our results suggest show that to be effective policies designed to deter corporate misconduct must account for interactions between external governance, such as laws and their enforcement, and internal governance mechanisms such as managerial compensation. The analysis offers insights into how to improve the deterrence of corporate misconduct.

1 Introduction

When corporations are created, they obtain many legal rights that enable them to own assets, sign contracts, and operate in markets. The only commitment they generally must make is to act lawfully. When a corporation breaks the law, it may face enforcement actions that result in sanctions such as monetary fines. Law enforcement might also be directed at individuals who break the law while acting on behalf of the corporation. This paper analyzes the interaction between prevailing law enforcement practices and corporate governance mechanisms such as managerial compensation structures. We show how these interactions weaken law enforcement and how some law enforcement practices can actually exacerbate corporate misconduct and harm to society. Our results have important policy implications.

In a well-known 1970 essay, Milton Friedman argued that "the social responsibility of managers is to make as much money as possible while conforming to the basic rules of the society, both those embodied in law and those embodied in ethical custom."[1] If managers break the rules of society

* Admati (admati@stanford.edu) and Pfleiderer (pfleider@stanford.edu) are at Stanford Graduate School of Business. Atkinson (natkinson@wisc.edu) is at University of Wisconsin Law School. We are grateful to Miriam Baer, Greg Buchak, Guiseppe Dari-Mattiacci, Peter DeMarzo, Dhammika Dharmapala, Michael Fishman, Scott Ganz, Louis Kaplow, Vic Khanna, Doron Levit, Amir Licht, Nadya Malenko, Ellen Podgor, Elizabeth Pollman, Jed Rakoff, Amit Seru, Roy Shapira, Chris Snyder, Jennifer Taub, Jason Yackee, and seminar participants at George Mason University, USC, Western Ontario, and participants at the 2024 Berkeley ESG conference, the 2025 American Law and Economics Association (NYU), and the 2025 NBER Law and Economics Summer Institute for valuable discussions. Sophie Molins provided helpful research assistance. The authors do not have any relevant financial interests or any conflicts of interests related to the research described in this paper.

[1] "The Social Responsibility of Business is to Increase its Profits," Milton Friedman. *The New York Times*, 1970.

in their quest to "make as much money as possible," corporations can cause enormous harm, which all too often involves deaths, physical and mental injuries, pollution, and massive financial losses. In such cases "making as much money as possible" can hardly be regarded as "socially responsible." The recent history of numerous corporate scandals that have caused much harm suggests that laws and their enforcement have not been effective in ensuring that corporations are socially responsible, which raises research and policy issues that merit more study by economists and others. This paper examines the economic forces related to law enforcement as practiced in the corporate context as they interact with standard corporate governance mechanisms.[2]

Much of the research on corporate governance is concerned with making sure that the interests of the directors and managers who make key decisions on behalf of the corporation are aligned with the interests of the shareholders. The focus on *internal governance* in the corporate governance literature is based on the implicit assumption that *external governance* mechanisms including law enforcement, which control the relations between the corporation and stakeholders such as customers, employees and others, work well enough to guarantee that corporations and those acting on their behalf comply with laws. In other words, the standard approach to corporate governance presumes that maximizing shareholder wealth "would best maximize corporate value and also social welfare, as other bodies of law could regulate corporate externalities" (Lund and Pollman (2021), p. 2574). Admati (2017, 2021) point out that if laws are ineffective, financialized corporate governance can harm society.[3] We show in this paper that ineffective law enforcement that fails to deter corporate misconduct is further undermined by corporate governance practices aimed at maximizing share prices, which can exacerbate social harm. Our analysis produces important new insights about how law enforcement can fail in the corporate context, and suggest potential improvement in policy.

Laws are enforced by governments and at times by people in the private sector who may have standing to take legal action. Since enforcement is generally costly, important economic trade-offs can arise. Starting with Becker (1968), many economists have explored optimal deterrence, assuming a set of costs and frictions and then weighing the social costs and benefits of different enforcement policies. Actual enforcement policies, however, are the result of complex political processes within and across various institutions.[4] The outcomes can involve either suboptimal over-enforcement or suboptimal under-enforcement.

There is a growing consensus that under-enforcement is common in the corporate context (see e.g. Garrett, 2014; Eisinger, 2017; Taub, 2020; Coffee Jr, 2020). Detecting law breaking by corporations and estimating the corresponding harm, is difficult.[5] Large, complex, and well-resourced corporations can make enforcement efforts by less well-resourced governments or by private par-

[2]Admati and Hellwig (2024), for example, focuses on poorly designed regulations that allowed the existence of an excessively fragile and dangerous financial system. The last chapter of the book, entitled "Above the Law?" discusses law enforcement related to the current paper.

[3]Relatedly, Zingales (2015) questions whether finance benefits society, pointing to the many cases of fraud. Admati and Hellwig (2024) elaborate on the issues in the context of banking.

[4]For example, there is a great deal of discretion given to agencies and prosecutors about what actions will be taken or whether any action is taken at all. Decisions may be taken based on limited resources, and are clearly subject to political forces (see, e.g., Podgor (2012) and Podgor (2018)).

[5]Corporate misconduct such as fraud and money laundering often persists without detection for years (Levi and Reuter, 2006). Dyck, Morse and Zingales (2010) and Dyck, Morse and Zingales (2024) estimate that only a third of frauds among large publicly traded firms are ever detected and that the total cost of fraud among these firms is $830 billion per year. Rudman et al. (2009) estimate the cost of fraud in the healthcare industry to be between $100 billion and $170 billion per year. The Economic Policy Institute estimated in 2017 that US workers lose over $8 billion per year to wage theft—the largest form of theft committed in the United States (https://www.epi.org/publication/employers-steal-billions-from-workers-paychecks-each-year/)

ties costly and difficult, weakening deterrence. Diffuse responsibility within corporations makes it difficult to identify individuals who should be held accountable for corporate misconduct. In addition, law enforcement is often ineffective because of authorities' concerns that innocent stakeholders will suffer collateral harms or because of other political considerations that may protect "important" corporations.[6] Finally, the limited liability of corporations, the tax subsidies that encourage the use of debt funding, and the option corporations have to seek bankruptcy protection and thus limit their liabilities (including from corporate misconduct) all serve to reduce the potential consequences of misconduct and the potential deterrent impact of law enforcement.[7]

When the profits a corporation gains from misconduct exceed the expected fines or other consequences that may result from the misconduct, the consequences may be seen as a "cost of doing business," and "good" business models may include profitable misconduct.[8] Since managers have financial incentives related to the ownership of stock and options and to performance-based bonuses, they will have incentives to engage in misconduct that increases corporate profits, unless they fear personal consequences for doing so. Profit-maximizing corporations in turn will find managers who are motivated by financial incentives and reward them in ways that ensure they act in the interest of the corporation, which can include engaging in misconduct. As we explore in this paper, managerial compensation can often be designed offset or reduce the effects of possible enforcement actions. [9]

To study the issues and explore the effectiveness of realistic features of law enforcement policies in deterring profitable corporate activities that create substantial social harm, we develop and analyze a model that captures key elements of corporate governance and enforcement practices. We maintain the standard principal-agent approach to corporate governance in which corporations set managerial compensation to align the interests of managers with those of shareholders. In our model, managers choose costly actions that vary in their impact on the corporation and on society. We examine how corporations modify their compensation structures to adjust to law enforcement policies when enforcement of the law falls short of what is necessary to deter misconduct effectively. We also consider a range of commonly-used enforcement policies aimed at either the corporation or at corporate managers and examine their effectiveness for reducing the social harm from profitable misconduct.

[6] On political economy forces and considerations involving collateral consequences, see Karpoff, Lee and Vendrzyk (1999), Atkinson (2023*a*) and Admati and Hellwig (2024). Enforcement policies are often determined by political actors. For example, as of early 2025, enforcement of the Foreign Corrupt Practices Act as well as the Corporate Transparency Act in the U.S, meant to combat bribery and money laundering respectively, has been "halted" or "paused."

[7] Admati et al. (2018) shows that indebted (already-leveraged) corporations have strong incentives to further increase their indebtedness, which can then limit the extent of subsequent fines (Atkinson, 2023*b*). On the use of bankruptcy to manage litigation and avoid accountability, see Jacoby (2024)

[8] Atkinson (2023*c*) estimates that the largest violations of the Clean Air Act are profitable for firms even after paying fines. Findley, Nielson and Sharman (2024) establish through experiments that compliance with anti-money laundering laws is often poor in developed nations, including the US, despite numerous nations committing to enforcing these laws. Byars (2025) points out that enforcement practices such as Deferred and Non-Prosecution Agreements can contribute to corporate recidivism. In a particularly egregious scheme dubbed *Cum-Ex*, which persisted for many years in Europe (and may still take place), numerous banks across Europe engaged in tax fraud that effectively robbed governments and their citizens of tens of billions of euros. Billions in ill-gotten profits have not yet been returned. On the fraud, see Chapter 17 of (Admati and Hellwig, 2024). Recent estimates put the total amount stolen of European citizens at about €150 billion.

[9] Efendi, Srivastava and Swanson (2007) shows that CEOs with in-the-money options are more likely to misreport earnings and Raghunandan (2021) finds evidence that some corporate executives use wage theft as a form of earnings management. Egan, Matvos and Seru (2019) show that financial advisors found to have committed misconduct find subsequent employment relatively easily. Desai and Dharmapala (2006) show a positive relationship between tax avoidance and high-powered managerial incentives.

We show that when law enforcement is weak, contracts that ensure that the manager has incentives to maximize shareholder value can lead the corporation to engage in profitable yet harmful activities. Of course, this result is not surprising—if the manager is aligned with the shareholders, the manager will pursue harmful activities when those are profitable. What is more interesting is how managerial contracts adjust to changes in the enforcement regime. Whereas for a *given* incentive contract, managers will engage in less misconduct if fines are increased, the same is not always true when corporations can adjust the contract as the standard principal-agent model would predict. As we show, these adjustments can undo much—or even all—of the harm-reducing effects of increasing fines. We consider various ways fines can be calibrated: to the level of ill-gotten profits, to measures of the amount of misconduct, or to the total harm caused. We then show how corporations adjust to these fines and other enforcement policies and how in some cases these adjustments can exacerbate the overall social harm.

Levying penalties directly on managers is seen by some as a better approach to law enforcement, since large fines imposed on the corporation may lead to collateral consequences for other innocent stakeholders. Aside from the fact that it may be difficult to impose liability on individuals in a complex organization, our analysis highlights the fact that corporations have ways to offset the deterrence effects of individual sanctions. We examine two ways the corporation can do so. First, the corporation can respond to increases in managerial liability by increasing the manager's equity stake, which can reduce or nullify the harm-reducing effects of the sanctions. Second, the ubiquitous presence of insurance and indemnification for corporate managers allows the corporation to limit the exposure of the manager to liability, again reducing deterrence and potentially exacerbating social harm from profitable misconduct (Baker and Griffith, 2019).

Because of the difficulties in detecting and sanctioning corporate misconduct in a timely manner, authorities attempt to encourage corporations to cooperate in identifying misconduct and assisting law enforcement. Of course, some misconduct by corporate agents involves individuals taking actions that harm the corporation (e.g., embezzlement). Other actions that can be taken by employees harm society but have little or no effect on the corporation except through fines and other sanctions imposed on the corporation (e.g., an employee drives negligently while doing company business). When employees act in ways that produce no benefits but only costs to the corporation, the corporation and its shareholders have incentives to search out and punish these employees to deter this type of bad behavior. In all these cases the interests of the corporation and law enforcement are aligned. We examine cases where the corporation is conflicted with law enforcement because misconduct benefits the corporation and shareholders. This creates incentives to encourage rather than punish or reveal the misconduct.

We explore the effectiveness of two types of programs designed to encourage corporations to cooperate with authorities in exchange for discounted fines. The first type we study is a "compliance program" where the firm voluntarily implements procedures that enhance monitoring and increase the probability of detection. The other type rewards the corporation for stepping forward and self-reporting misconduct. One intended benefit of these programs is they lead to earlier detection of misconduct, which may reduce the amount of social harm.

For each of these programs, we show that although they generally reduce the duration of misconduct, they can also make matters overall worse. By offering discounts for cooperation, effective penalties are lowered, which makes misconduct more profitable. This fact means that the corporation will pursue the harmful activity with greater intensity up to the point of detection or profit-maximized self-reporting. This greater intensity can ultimately create more harm even though the misconduct will likely be detected earlier. Since programs offering discounts in fines can be counterproductive, a better approach may involve making compliance programs specific

and mandatory without giving fine reductions. Other steps taken to increase the likelihood of detection (for example by encouraging and providing protection for whistleblowers) are less likely to be counterproductive. Finally, increasing the level of fines will generally be a more effective and reliable way to reduce harm.

Our paper speaks to controversies related to corporate purpose and its interaction with corporate law. Whereas the standard approach to corporate governance essentially equates corporate purpose with the creation of "shareholder value," the CEOs of 181 US corporations committed their firms to benefit of *all* stakeholders, including customers, employees, suppliers and communities as well as shareholders. Claims that corporations should have a broader purpose are related to the growth of Environmental, Social, and Governance (ESG) metrics, which purportedly enable investors to monitor corporations' impact across a range of socially relevant issues. The very existence of these concerns about purpose and how corporations affect other stakeholders seems to reflect a widespread recognition that governments are failing to set and enforce proper rules to ensure that corporations act in society's interests. This recognition seems to have led to attempts to make investors and others more responsible for ensuring that corporations account for how they affect stakeholders and social welfare.

Whether pressuring corporations to meet broader corporate objectives will improve overall welfare has been the subject of much debate.[10] For our purposes, however, we note that no matter how corporate law and key decision makers within corporations define and operationalize corporate purpose, *corporations must comply with all laws.* Once realistic law enforcement frictions are then incorporated into the analysis, we show that what might be considered "good" corporate governance practices, ones that succeed in aligning the interests of shareholders and managers, might produce bad outcomes for society.

In addition to relating to research on corporate governance and corporate misconduct discussed above, our paper also relates to the optimal deterrence literature, which examines how to structure legal sanctions optimally to deter wrongdoing. Particularly relevant to our analysis are discussions of the choice of whether to impose liability on the manager or the corporation, the choice between strict versus duty-based liability regimes, and the choice between public and private enforcement of the law.[11] We must emphasize that our analysis does not involve positing a set of frictions and then identifying the optimal enforcement policies for deterrence in the presence of those frictions. Instead, we start by assuming that enforcement policies are weak, an assumption that is strongly supported by the evidence, and we examine how internal governance decisions respond to changes in enforcement policies in weak enforcement environments. Our findings reveal some important considerations for improving enforcement and deterrence, but we are not deriving specific policies that we claim are optimal.

[10] Hart and Zingales (2017) argue that shareholder value is not the same as shareholder wealth and that corporations should attempt to aggregate over shareholder preferences. However, this aggregation is challenging (Döttling et al., 2024). Other complications in implementing broader corporate purpose include measurement issues (Raghunandan, 2021; Baker et al., 2024), legal barriers related to corporate law and other constraints on investors (Atkinson, 2020; Badawi and Partnoy, 2022), and, more recently, political backlash. Some jurisdictions require disclosures about, e.g., emission metrics. The Companies Act, 2013 in India also mandates such things as the formation of Corporate Social Responsibility (CSR) committees for eligible companies and mandatory CSR spending (at least 2% of average net profits over the previous three years for qualifying companies).

[11] Becker (1968) and Polinsky and Shavell (2007) provide general frameworks for the optimal deterrence literature. On the choice between corporate and managerial liability, see Kornhauser (1982), Polinsky and Shavell (1993), Argenton and van Damme (2014), and Deng, Goldman and Kahl (2021). On strict versus duty based liability, see Newman and Wright (1990), Arlen (1994), and Arlen and Kraakman (1997). See Polinsky (1980), DeMarzo, Fishman and Hagerty (2005), and McAfee, Mialon and Mialon (2008) for the reasons that private enforcement or enforcement through self-regulatory organizations cannot fully substitute for public enforcement.

Our paper is also related to the literature on the interaction of managerial compensation and corporate tax avoidance. Chen and Chu (2005) investigate (profitable) tax evasion in a principal-agent model.[12] Crocker and Slemrod (2005) similarly analyze contracts when profitable tax evasion is possible, and analyze how those contracts change with changes to the enforcement regime. Crocker and Slemrod show that fines on the manager are generally more effective for deterrence than fines on the corporation. Similar to our results, they show that adjustments to managerial compensation can partially offset and undermine the effect of enforcement efforts.

The remainder of the paper is organized as follows. In the next subsection we provide some institutional background on corporate law enforcement along with some illustrative examples showing how enforcement can be weak and lead to inadequate deterrence. In Section 2 we develop our formal model. Section 3 shows various ways that internal governance mechanisms can reduce or even nullify the deterrent effects of fines. It also shows how basing fines on harm caused by misconduct rather than profits that are gained can lead to more harm. Section 4 considers the role that compliance programs and self-reporting might play in deterring corporate misconduct and shows that cooperation motivated by discounts in fines can actually increase harm and reduce welfare. Section 5 offers policy recommendations and concluding remarks.

1.1 Institutional Background and Illustrative Examples

There are many laws that place restrictions on what corporations can do and establish requirements for what they must do.[13] In the US, corporations can be held criminally liable for breaking certain laws, including for any crime committed by any of their employees in the scope of employment.[14] Laws are accompanied by enforcement mechanisms that specify who might take action, as well as the process of adjudication and any consequences if a corporation breaks the law. In most cases the consequences of enforcement actions brought against corporations are determined by out-of-court settlements, which include Deferred Prosecution Agreements (DPAs), Non-Prosecution Agreements (NPAs), and civil settlements.[15]

Corporations obviously cannot be incarcerated. The maximum criminal and civil penalties for a corporation may also be quite small, as illustrated by the case of PG&E' criminal cases (discussed below) and by restrictions on corporate penalties in Germany and elsewhere. Given that fines are often formulaic, they often fail to account reliably for either the social harm or the private

[12]Relatedly, Desai, Dyck and Zingales (2004) study how tax enforcement may affect manager's propensity to expropriate income. Khan, Srinivasan and Tan (2017) find significantly positive relation between institutional ownership, thought to provide incentives for managers to achieve financial performance, and corporate income tax avoidance.

[13]Corporate law, under which corporations are created, focuses on the rights and responsibilities of directors, executives and shareholders. Securities laws aim to protect investors and in doing so help corporations raise funding. The legal environment in every jurisdiction in which a corporation operates includes numerous other rules. These include those concerning property, contract, bankruptcy, antitrust, torts, tax, labor, environmental, consumer protection, civil rights, anti-discrimination, anti-bribery, anti-money laundering, human rights, and speech. Some sectors such as banking, utilities, defense, food and drugs, aviation, mining, housing and media, also have specialized rules to ensure safety and to protect the public.

[14]This "vicarious liability" refers to the doctrine of respondeat superior, see Thomas (2021) for a discussion of the issues controversies related to corporate criminal liability and the vicarious liability doctrine. In other jurisdictions (most notably Germany), it is essentially impossible to charge corporations of crimes.

[15]Two sources on corporate misconduct and their legal consequences for corporations are the Corporate Prosecution Registry (https://corporate-prosecution-registry.com/) and the Corporate Violation Tracker (https://violationtracker.goodjobsfirst.org/). The use of DPAs and NPAs in the US, which has increased dramatically in the last two decades, has been criticized as excessive, for example by Garrett (2014, 2020); Taub (2020), Rakoff (2021), and Byars (2025).

benefits the corporation gained from the misconduct (Alexander and Cohen, 2015). As discussed earlier, authorities may significantly reduce fines and other sanctions on corporations because of concerns about the potential collateral consequences or for other reasons. In addition, fines may be insufficiently adjusted to account for the time value of money given the often substantial lag between the occurrence of misconduct and the payments of fines (Atkinson, 2023*c*).

As discussed in the introduction, fines and other sanctions may be substantially reduced for a corporation charged with misconduct if the corporation had established a compliance program and is otherwise seen to be cooperating with authorities. Both the US Sentencing Guidelines and Department of Justice Policy call for substantial reductions for corporations that self-report or have effective compliance programs. Cooperation with authorities can also enable corporations to avoid prosecution or having to plead guilty to misconduct, thus avoiding costly private litigation and reducing substantially the costs to the corporation of misconduct.

Providing incentives for corporations to help law enforcement through voluntary compliance and self-reporting is a common approach by the US Department of Justice (DOJ). In a May 2025 speech, for example, the Head of the Criminal Division emphasized the generous benefits companies can obtain from voluntary self- reporting, and specifically promised that such cooperation will lead in most cases to the DOJ declining to file any charges. Related practices were recently announced by the Serious Fraud Office in the UK. As mentioned, our analysis shows that one cannot assume that these programs necessarily help reduce the harm to society from corporate misconduct. The voluntary nature of this cooperation can backfire and actually increase harm when misconduct is profitable.

Our model is mostly aimed at capturing the effects of law enforcement by public authorities. In some cases private actors, including victims of corporate misconduct and shareholders who claim to have suffered losses, can sue the corporation and its officers and directors. Shareholders can sue in a number of ways. For example, they can sue under US securities laws that are meant to ensure proper disclosures. Such lawsuits are viewed by many as highly inefficient and not an effective way to deter securities fraud (see e.g., Bratton and Wachter, 2011). Corporate law also allows shareholders to sue directors and officers for failure of oversight, but shareholders have little incentive to file such lawsuits in cases where misconduct is profitable and the shareholders actually benefit. Lawsuits by shareholders against directors and officers are often dismissed if it is established that the corporation had a "sufficiently" effective compliance program (Shapira, 2024).

Whereas lawsuits can be brought against the corporation and its officers by other stakeholders, the right to take legal action against corporations is often limited. For example, customers and employees routinely sign away their legal rights and agree to use opaque private arbitration processes to settle claims against corporations (Estlund, 2017). The US Supreme Court has also limited corporate liability in a variety of ways (Pollman, 2021). Even when private enforcement actions are deemed successful, misconduct may still be profitable (see, e.g., Shapira and Zingales, 2023). Overall, private enforcement actions generally have quite limited value in the deterrence of corporate misconduct.

For the type of profitable misconduct this paper focuses on, it is exceedingly rare for individuals, particularly for directors and executives, to face meaningful criminal liability. Buell (2024) examines recent cases and shows that the government has been largely unsuccessful in most of these cases, partly because of the high bar necessary to show personal intent to commit crimes. Eisinger (2017) and Rakoff (2021) conjecture that the focus on corporate settlements and the reluctance to bring cases against individuals since the early 2000 are likely related to constrained resources and to the career incentives of prosecutors.

Managers also tend to be shielded from civil liability in cases of corporate misconduct. Given the way laws are written and given the ubiquitous use of insurance and indemnification, officers and directors, who are the key decision makers, rarely face meaningful consequences for corporate misconduct (Black, Cheffins and Klausner, 2005; Baker and Griffith, 2019).

Another possibility in cases where misconduct and harm are particularly salient is that corporations, or their executives and board members, incur reputational losses that affect their profits or future opportunities. Given the frequency at which corporate misconduct occurs, the deterrent effects of reputational losses appear to be limited.

Some of the key issues around law enforcement and corporate misconduct are illustrated through the three well-known cases of Pacific Gas and Electric, Boeing, and Volkswagen, corporations that broke laws and caused great harm. We discuss these cases briefly below. Similar elements can be seen in many other examples, such as General Motors, HSBC, Purdue Pharma, many banks involved in Cum-Ex tax fraud, and others.

Pacific Gas and Electric (PG&E) is a publicly traded Northern California Utility.[16] Following a 2010 explosion of a gas pipeline that killed eight people and injured 58, a jury found PG&E guilty of five violations of the Natural Gas Pipeline Safety Act in late 2016. The corporation paid only a $3 million fine. The sentence required PG&E to maintain a monitor for five years of probation starting in January 2017.

In 2017 and 2018, wildfires caused enormous harm and many deaths throughout California. In June 2020, PG&E pleaded guilty to 84 counts of manslaughter related to criminal negligence caused by its poorly maintained equipment that played a major role in the 84 deaths associated with the 2018 fire. The company paid only $3.4 million ($41,667 per lost life), which amazingly was the maximum penalty that could be imposed. The presiding judge noted that a natural person found guilty of 84 manslaughter charges would normally be sentenced to 90 or more years in jail. No individual associated with PG&E was charged with any crimes either in the context of the 2010 explosions or related to the 2018 fire.

When PG&E's monitorship for the federal crimes related to the 2010 explosion ended in January 2022, the federal judge overseeing PG&E expressed strong criticism of the company's behavior, noting that during the five-year probation, PG&E was responsible for multiple catastrophic wildfires, leading to significant loss of life and property damage. The judge described PG&E's actions as a "crime spree." The company continued to face litigation, agreeing in April 2022 to pay more than $55 million to avoid criminal prosecution for two major wildfires sparked by its aging power lines in 2019. A civil lawsuit against former PG&E executives and board members in 2022 was settled for $117 million, fully covered by the Directors and Officers liability insurance paid for by PG&E.

PG&E was able to reduce its legal liabilities towards victims through Chapter 11 bankruptcy, while shareholders retained positive equity. This case illustrates how bankruptcy can be used strategically as a way for corporations to to manage litigation risk, as well as how it can lead to inefficient and unjust outcomes for victims of misconduct or tort harm.Jacoby (2024) discusses the problems with how US bankruptcy process and its implementation treat victims, and Admati (2025) mentions the case of PG&E, described in more detail in Blunt (2022).

Atkinson (2023*b*) Another revealing example involves safety issues at Boeing, a US corporation long considered a "national champion." In selling its 737 Max planes Boeing made false statements to the Federal Aviation Authority and to purchasing airlines about safety and the training necessary for pilots to fly the new plane. Two plane crashes, one in 2018 and the other a few months later in 2019, resulted in 346 deaths. After the first crash, Boeing's CEO asserted

[16]Blunt (2022) gives more details on the history PG&E.

that the plane was safe, and it continued to operate for some time in the US and other countries (Robison, 2022).

In 2021, the Department of Justice (DOJ) settled charges with Boeing in a Deferred Prosecution Agreement (DPA). Whereas the headline figure was $2.5 billion, less than 10% of the total amount was a criminal penalty. The rest of the settlement consisted of $1.77 billion to be paid to airline companies that had bought the planes and $500 million to establish a fund for crash victims. The settlement, which was coupled with Boeing's promise to improve its compliance program, has been criticized by legal experts (see e.g. Coffee, 2022). However, Boeing continued to be plagued with safety issues and after the DOJ determined that it violated its promises under the settlement agreement, Boeing pleaded guilty in July 2024 to a felony charge of conspiring to defraud the federal government, and it agreed to pay $487.2 million and to further strengthen its compliance and safety programs. However, a judge rejected the deal in December 2024, criticizing some elements of it. In May, 2025 the DOJ announced a Non-Prosecution Agreement that did not include a felony guilty plea.

Boeing received "cooperation credits" when negotiating penalties with DOJ both on the basis of promising better compliance and by also promising to cooperate with DOJ regarding potential law violations by its employees. It ostensibly made an effort to meet the latter commitment by providing information that led to criminal charges against former test pilot Mark Forkner on four counts of wire fraud. A jury in federal district court in Texas deliberated less than two hours before finding Mr. Forkner not guilty on any of the counts. Had the DOJ relied on its own investigation to find individuals who may have been responsible for Boeing's misconduct, it may have been able to hold higher level executives accountable.[17]

As for Boeing's directors, a shareholder lawsuit against them related to Boeing's oversight of the 737 MAX's design and development was settled for $237.5 million. However, the company took pains to note that "the settlement will be entirely paid by insurers rather than by Boeing or by its officers and directors."

In the so-called Dieselgate scandal, Volkswagen, the German car manufacturer, used special sensing equipment and software (called the "defeat device") to fraudulently manipulate the results of emissions tests on vehicles that it sought to market as "clean diesel." These deceptions resulted in VW vehicles emitting 40 times the legal limit of nitrogen oxides during normal driving, causing smog, respiratory disease and premature deaths. Volkswagen started using the "defeat device" in 2006 and over the years it was fraudulently used on many millions of cars of different models sold in many counties. It was not until 2014 that a team of researchers at West Virginia University uncovered evidence that created suspicions about Volkswagen meeting the clean air standards. As authorities started investigating, Volkswagen did not come forward but instead denied the allegations. Its executives destroyed evidence and sought to cover up the misconduct, and they tried for a long time to falsely put the blame a small group of low-level employees (Ewing, 2017).

Enforcement actions against Volkswagen and some of its executives and other agents started in the US in 2015, but reactions to the scandal in Europe were very slow to emerge. The severity of the penalties and sanctions has varied greatly across jurisdictions.[18]. Some cases have not been resolved, including against the former CEO. In some cases charges were brought against relatively low level employees considered by some to be "scapegoats." At the same time, Volkswagen shielded

[17]For an extensive (and highly critical) discussion of the Deferred Prosecution Agreement in the Boeing case, and of the case against Mr. Folkner, see Coffee (2022). On the Yates memo that directs DOJ to seek individuals and promises credits for corporations who provide information about individuals, its impact and the potential problems it raises, see Oded (2016), Copeland (2014), Copeland (2016) and Garrett (2020).

[18]For more discussion of the consequences in the first few years, see Ewing (2017) and Welch Jr (2019).

two executives by paying €9 million (about $10 million) to settle *criminal charges* against them for market manipulation related to the scandal. It later collected €270 million from their Directors and Officers Insurers to cover expenses associated with their legal problems.. To this day, millions of diesel cars emitting dangerous levels of nitrogen oxides continue to be on the road in Europe.

2 Model

We formally analyze the strategic interactions among three key actors, the corporation, its manager, and the government, which taken together determine corporate activities that affect societal welfare. The government sets external enforcement policies, including fines for harmful corporate activities that can be levied against the corporation and its manager. The corporation designs the compensation contract for its manager to maximize shareholder profits net of any fines paid by the corporation. The manager chooses levels of corporate activities based on the compensation contract and personal cost of effort.

1. **Corporate Activities.** The corporation has two activities available, one that on net benefits society and the other which on net harms society. Its manager chooses the level of each activity, where x_b is the level chosen for the beneficial activity and x_h is the level set for the harmful activity. Because of resource constraints the activity levels must be chosen so that $x_b + x_h \leq 1$.

2. **Law Enforcement.** The government specifies a policy for imposing fines that in general will depend on x_h, the level of the harmful activity. The fine might be imposed on the corporation or the manager. In Section 3 we assume that the government is able to detect that $x_h > 0$ with certainty. (We relax this assumption in Section 4 when we discuss compliance programs and self-reporting that affect the probability of detection.)

3. **Managerial Contract.** The corporation offers a compensation contract to the manager that has the form

$$Comp = S + \alpha \times \Pi_c, \tag{1}$$

where $S \geq 0$ is a fixed salary, $\alpha \in [0, 1]$ is the profit-sharing parameter, and Π_c is corporate profits. The corporation chooses S and α to maximize its expected profits, taking into account the enforcement policy.[19]

4. **Corporate Profits** The corporation seeks to maximize its net profit, which equals

$$\Pi_c(x_b, x_h) = \pi_b\sqrt{x_b} + \pi_h\sqrt{x_h} - Comp - Fine_c, \tag{2}$$

where $\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}$ is the component of profits earned from the corporation's activities with $\pi_h > 0$ and $\pi_b > 0$ and $Fine_c$ is the fine (if any) levied against the corporation. Note that the corporation confronts decreasing returns when it allocates more resources to either activity.

[19] Our assumption about the contract is based on the common practice of providing stock-based compensation to align managers with shareholders who care about the share price. In our setting, x_h captures misconduct that is often illegal, and contracts that specify taking an illegal action are not enforceable in court. Contracting on x_h may also subject the manager to high risk of civil lawsuits by tying him directly to the harmful action.

5. **Manager's Choice of Activity Levels.** After accepting the contract, the manager chooses activity levels (x_b, x_h) to maximize his own utility, which is given by

$$U_M(x_b, x_h) = Comp - \beta \times (x_b + x_h) - Fine_m, \tag{3}$$

where $\beta \times (x_b + x_h)$ represents the disutility of effort with $\beta > 0$ and $Fine_m$ is the fine (if any) imposed on the manager based on the government's policies.[20] If β is very large, the shareholders will find it too expensive to compensate the manager in a way that maximizes profits. This will not occur if

$$\beta < \frac{\sqrt{\pi_b^2 + \pi_h^2}}{4}.$$

We will make the realistic assumption that this condition holds. As we will show, when it is violated, the shareholders would need to share more than 50% of the profits with the manager, which is rarely, if ever, seen in practice in large corporations.

6. **Impact on Social Welfare:**

The impact on social welfare caused by the corporation is equal to

$$W(x_b, x_h) = b\sqrt{x_b} - h\sqrt{x_h}, \tag{4}$$

where both b and h are positive. We will define B and H to be the total effect on welfare caused by the two activities, i.e., $B = b\sqrt{x_b}$ and $H = h\sqrt{x_h}$. Harms and benefits are measured on a net basis, so that H should be thought of as *all* of the harms caused by the harmful activity *minus* any benefits it might produce. This means that to determine the total effect the harmful activity has on welfare, we do not need to subtract $\pi_h\sqrt{x_h}$ from H, since H is already net of any benefits that might be related to profits.[21] For simplicity we do not include the manager's disutility of effort in $W(x_b, x_h)$ since it is assumed to be very small relative to the other benefits and harms.[22]

This setup provides the foundation for analyzing how changes in the government's enforcement policy and the corporation's internal governance, *Comp*, interact, particularly when external enforcement is weak.

[20] Note that both $Fine_c$ and $Fine_m$ can be interpreted as the expected value of the fine. In section 4 we will explicitly consider a multi-period model where the probability of detecting the corporation's misconduct in any period is less than certain. This makes the fines levied uncertain and decisions are made based on expectations.

[21] One should not assume that corporate profits should always be counted as a benefit. Consider a pharmaceutical company that markets an opioid drug in a way that thousands of people who should not take the drug become addicted (this is the harmful activity). If, when accounting for overall effects of the harmful activity on social welfare, one wants to include the corporation's profits from this activity as a benefit, one must recognize that this is more than offset by the fact that these profits are the result of customers paying for a drug they should not be taking. For this reason profits are not a benefit that should be deducted from all of the harms that occur when people become addicted to opioids.

[22] Since the disutility of effort is equal to $\beta \times (x_b + x_h)$ it only depends on the sum of the two activity levels. When resources are fully deployed, as will be optimal, the total disutility of effort will not vary and for this is another reason it can be ignored.

3 Managerial Compensation and Misconduct

We begin in subsection 3.1 examining the baseline case of no external enforcement and then analyze how corporations respond to fines imposed on them (subsection 3.2), or on their managers (subsection 3.3). Throughout our analysis and in keeping with what the evidence shows is quite common, we assume that the fines imposed are in expectation less than the profit gained from the harmful activity. Clearly, if the expected fine exceeds the profit from the harmful activity, the corporation would have no incentive to engage in misconduct and the corporation only takes socially beneficial actions. We will examine three ways that fines might be determined. First, a fixed fine might be imposed if the level of the harmful activity exceeds a given level. Second, a variable fine might be imposed based on the level of profits the firm gains from the harmful activity. In subsections 3.2 and 3.3 we explore what happens when these types of fines are levied against the corporation or the manager respectively. A third type of fine would be calibrated to the total harm caused by the misconduct, which might create incentives for the corporation to reduce the rate of harm caused by the harmful activity. We explore this case in subsection 3.4.

3.1 No External Enforcement

We first assume that there is no external enforcement: that is, setting $x_h > 0$ is not penalized by any fines even though it is harmful to society. Since $h > 0$, society is necessarily worse off whenever the corporation sets $x_h > 0$.

If the manager is paid only a fixed salary, he clearly maximizes his utility by choosing $x_b = x_h = 0$. The corporation solves this basic agency problem involving effort by making compensation depend on profits.

Proposition 1. *The manager will have the incentive to choose the levels of x_b and x_h that maximize the total profit attainable for the two activities if compensation is set optimally equal to a constant plus a fraction of the corporation's profits.*

Proof: See appendix.

Absent external enforcement, the profit-maximizing levels of both activities are positive. Proposition 1 therefore underscores the challenge at the heart of our analysis: "good" internal governance, i.e., internal governance that fully aligns the manager with the corporation, ignores the social welfare effects of the harmful activity. Specifically, whenever the marginal profits of the harmful activity are greater than those of the beneficial activity, the compensation contract gives the manager incentives to push more resources into the harmful activity. With no external governance, optimally aligning the interests of managers with shareholders can amplify the harm to society that the corporation creates. Note that since S, the salary component of the manager's compensation, is a constant, it will not affect the optimal choices of x_b and x_h. For that reason we will from now on ignore the salary component in most of our analysis. We will only mention salary when we discuss situations where the corporation might need to increase S to satisfy the manager's participation constraint given changes in enforcement and the manager's outside options.

The results of Proposition 1 are illustrated in the Numerical Example below. We will use this example and its parameters to illustrate most of our results.

> **Numerical Example.** Consider a simple numerical parametrization where $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $h = 15$. In this case the profit-maximizing equity share

is $\alpha = 0.004$, which leads to the optimal activity levels of $x_b^* = 0.64$ and $x_h^* = 0.36$. Corporate profits are $\Pi_c = 5$, and the manager's take home pay is $(0.004)(5) = 0.02$. This compensation more than offsets his disutility of effort, which is $\beta \times (x_b^* + x_h^*) = 0.01 \times (0.64 + 0.36) = 0.01$.

From a social welfare perspective, the harm caused to third parties by the harmful activity is $H = h\sqrt{x_h} = 15\sqrt{0.36} = 9$. This is much greater than the profits the corporation makes by engaging in this activity, which are $\pi_h\sqrt{x_h} = 3\sqrt{0.36} = 1.8$, and indeed is greater than total corporate profits of 5.

3.2 Internal Governance Responses to Corporate-Level Fines

There is a simple and obvious way to ensure that the corporation only takes beneficial actions. If the government imposes a fine that is greater than the benefit to the corporation from the harmful activity, then the activity will be fully deterred, and the corporation will not harm society. However, as was discussed above and as ample evidence shows, there are many cases where fines are set below the level that leads to complete deterrence. Our focus in this section is how internal corporate governance changes in response to suboptimal deterrence when fines are imposed on the corporation.

Suppose the government may fine the corporation and that the size of the fine may depend on the level of x_h (which for now we assume is observable). We will consider two types of fines. One is a fixed fine that is imposed whenever the corporation sets $x_h > \psi$, where ψ is some threshold greater than or equal to zero. The other is a fine that increases in the level of the profits the corporation gains from the harmful activity. We will show that there are major differences between fixed and variable fines. Whereas a fine that varies with the level of profits that the corporation gains from the harmful activity will generally reduce the harmful activity even if it does not fully deter it, a fixed fine may affect only profits and have no effect in reducing harm. It may therefore simply be seen as a cost of doing business.

First, consider the fixed fine:

Proposition 2. *If a fixed fine of F_c is imposed on the corporation whenever $x_h > \psi$, either the corporation sets the level of the harmful activity equal to ψ or sets it equal to the profit-maximizing level it would choose with no fine. In every case there is a compensation contract that gives the manager the incentive to do what is optimal for the corporation.*

Proof: See appendix.

Proposition 2 shows that a fixed fine can be completely ineffective in reducing the level of the harmful activity. In cases where the threshold ψ is higher than the level of x_h that is optimal for the corporation, the fine has no effect on the misconduct decision. In other cases, the corporation simply treats the fine as a cost of doing business and does nothing to change the level of the harmful activity. Only in cases where the fine is sufficiently high does it have an effect in reducing harm. Clearly, it would be optimal from a social welfare perspective to set the fine very high and set ψ equal to zero, but we know that the actual practice generally falls far short of that.

We now turn to the case where the fine is proportional to the corporation's profits derived from the harmful activity and the harm that is done. We assume that the corporate fine is equal to $f_c \times \sqrt{x_h}$.[23]

[23] Another interpretation of this fine is that it is fixed at some level as in the first case but the probability of the fine being imposed is linearly related to the level of profits the corporation realizes.

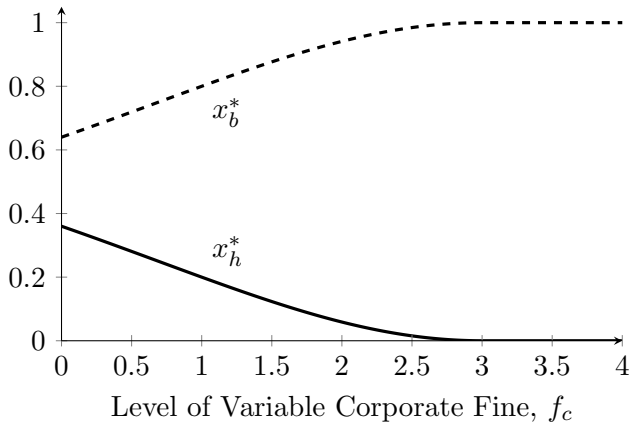


Figure 1: **Choice of activity levels as a function of fine rate.** This figure shows the choice of the beneficial and harmful activities (taking into account the change in the manager's compensation) as a function of a variable fine rate for the parametrization $\pi_b = 4$, $\pi_h = 3$, and $\beta = 0.01$.

Proposition 3. *If the fine is equal to $f_c \times \sqrt{x_h}$, the scale of the harm will be reduced if $0 < f_c < \pi_h$, and the harm will be eliminated if $f_c \geq \pi_h$. The manager can be motivated to make decisions to maximize profits net of fines with a simple change in the compensation contract.*

Proof: See appendix.

Whenever f_c is less than π_h, fines are viewed as a cost of doing business and the corporation chooses activity levels with a positive level of the harmful activity.[24]

$$x_b^* = \frac{\pi_b^2}{\pi_b^2 + (\pi_h - f_c)^2} < 1; \quad x_h^* = \frac{(\pi_h - f_c)^2}{\pi_b^2 + (\pi_h - f_c)^2} > 0. \tag{5}$$

These results can be illustrated in the context of the parametrization given in our Numerical Example above, as shown in Figure 1 and Figure 2. In Figure 1 we see that as the fine is increased from zero, the corporation will cut back on the harmful activity and increase the beneficial activity since the fine reduces the profitability of misconduct. But only when the fine equals or exceeds the profitability of the harmful activity ($f_c \geq \pi_h$) will the corporation desist. Figure 2 shows the effects of the fine on total profits and the corporation's impact on social welfare.

Importantly, in both the case of fixed fines (Proposition 2) and variable fines (Proposition 3), the corporation is adjusting managerial compensation, α, when fines change in order to induce the manager to choose the profit-maximizing effort levels. An increased fine decreases the corporation's profits for a given level of effort, and thereby decreases the manager's equity-based pay. Because the fine is effectively a tax on the rewards to the manager's effort, he will generally reduce effort in response to an increase in the fine level. The corporation then needs to increase the manager's equity compensation in order to compensate for this tax.

To see this point more clearly, recall that given the parametrization in our Numerical Example, the corporation set the manager's equity based pay at $\alpha = 0.004$. This is captured in Row (B)

[24]Recall that the rate of social harm h is measured *net* of any benefits that might be related to profits. This means that a fine that fully accounts for harm will necessarily be greater than π_h when $h > 0$.

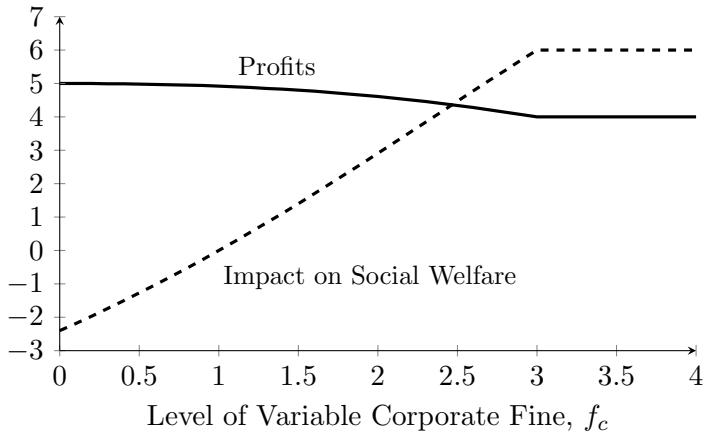


Figure 2: **Profits and Effect on Social Welfare as a function of fine rate.** This figure shows the equilibrium level of profits and effect on social welfare as a function of a variable fine rate for the parametrization $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $h = 15$.

of Table 1 below. Now consider the introduction of a variable fine, $f_c = 0.4$. If the corporation leaves the manager's contract unchanged, then the outcome is captured in Row (C). The fine lowers the profitability of the harmful activity, and therefore decreases the manager's equity-based compensation. Given the "tax" of the fine, the manager's privately optimal response is to reduce the harmful activity from 0.36 to 0.26. This in turn reduces corporate profits and social harm. However, now that there is a fine in place, setting equity based pay at $\alpha = 0.004$ is too low to maximize profits, because the corporation is no longer compensating the manager enough for the disutility of effort. By increasing the manager's equity based pay to 0.0042 (Row (D)), the manager once again chooses the profit-maximizing level of effort. Doing so increases corporate profits *and* increases social harm.

Case	**Equity Given to Manager**	**Beneficial Activity**	**Misconduct**	**Manager's Pay**	**Profits**	**H**
(A) No Fine	0%	0	0	0	0	0
(B) No Fine with Equity Adjustment	0.40%	0.64	0.36	0.02000	5.00	9.00
(C) Fine of 0.4; With No Equity Adjustment	0.40%	0.64	0.26	0.01820	4.53	7.65
(D) Fine of 0.4; With Equity Adjustment	0.42%	0.70	0.30	0.02003	4.77	8.17

Table 1: **The Effect of Variable Corporate Fines**. This table assumes the following numerical parametrization: $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $h = 15$. Each row illustrates an outcome for a given fine and pay structure. The last column shows H, the total harm created by the misconduct. Rows (A) and (C) are cases where the manager's pay is set sub-optimally, and Rows (B) and (D) illustrate the corporation's optimal equity adjustment to the given fine structure.

This point is important, because it illustrates the interaction between frictions in external enforcement and internal corporate governance considerations that affect managerial compensation. It is not simply the case that managerial compensation might incidentally give the manager

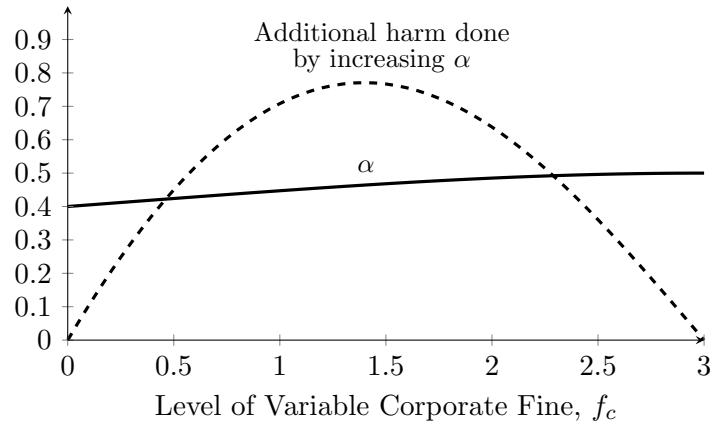


Figure 3: **Additional Harm as a function of increasing the profit percentage.** Updated optimal equity-based pay and resulting additional harm from fine-induced compensation adjustments for the parametrization $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $h = 15$ for variable fines.

incentives to engage in misconduct. Rather, when external enforcement is inadequate, the managerial contract is adjusted to reduce the impact of the fine, *making external enforcement even less effective than it might otherwise be.*

In both the case of the fixed fine and the variable fine, the optimal equity-based compensation is weakly increasing in the fine. For the case where the fine is equal to $f_c \times \sqrt{x_h}$, Figure 3 shows the additional harm created by this upward adjustment of α in the context of the parametrization from our Numerical Example. Figure 3 shows that as the fine increases from 0 to π_h, the corporation adjusts α upward from 0.4% to 0.5% to address the internal governance problem. This adjustment results in more harm than would be the case had the corporation kept α at 0.4%, the optimal value when no fine is imposed. In these situations internal governance actions that align the manager's interest with the shareholders clearly work against the interests of society at large.

3.3 Internal Governance Responses to Managerial Fines

As discussed above, imposing a sufficiently high fine on the corporation so as to make the misconduct unprofitable may not be possible or even desirable *ex post*. In those cases, the government may have another tool at its disposal, namely imposing penalties directly on the manager. Indeed, as we have discussed, the Department of Justice has repeatedly emphasized their intentions to increase individual-level liability for corporate misconduct.

We now consider the effectiveness of fines imposed directly on the manager rather than on the corporation. As we did for corporate fines, we will consider both fixed fines and fines that vary with the level of harmful activity. To keep things simple, for this section when specifying the compensation of the manager, we will ignore the constant term S, since this has no material effect on the analysis. As before it may be necessary to adjust S to make sure that the manager's participation constraint is satisfied. We restrict our attention to cases where fines are imposed only on the manager, not on the corporation and the manager.[25]

[25]Several optimal deterrence papers have examined how to divide liability between the corporation and the manager (Polinsky and Shavell, 1993; Argenton and van Damme, 2014; Deng, Goldman and Kahl, 2021). In our model, imposing liability on both the manager and corporation would generally increase deterrence. However, the results of this section coupled with the previous section show the steps that corporations could take to modify the manager's incentive contract may mean that the deterrence achieved by leveling a fine on both the manager and a

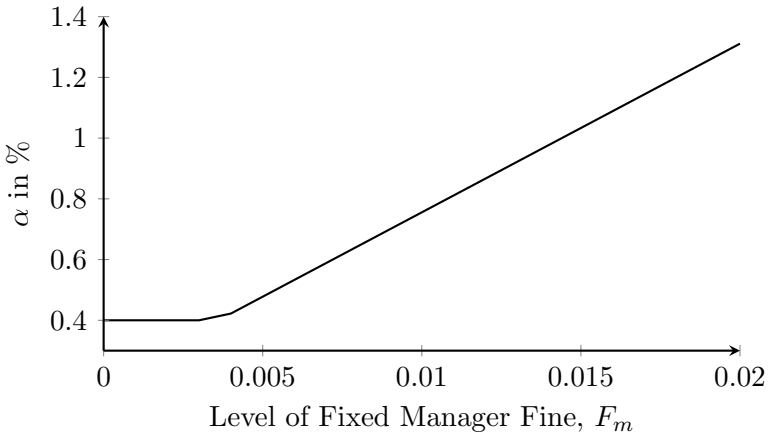


Figure 4: **Optimal Compensation as a Function of a Fixed Manager-Level Fine**. This plots the optimal managerial compensation as a function of a fixed-level managerial fine for the parametrization $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $h = 15$.

We start with the fixed fine. The following proposition shows that unless the fine is quite large, it will do nothing to reduce the level of the harmful activity.

Proposition 4. *Assume that a fixed fine of F_m is directly imposed on the manager whenever x_h is set above ψ. Let $\bar{\alpha} = 2\beta/\sqrt{\pi_b^2 + \pi_h^2}$ be the level of α that maximizes profits when there is no fine. There are three possibilities when the fixed fine is imposed:*

(1) The fine will have no effect on the choice of x_h with $\alpha = \bar{\alpha}$;

(2) The corporation can nullify the effect of the fine on the choice of x_h by increasing α above $\bar{\alpha}$;

(3) The fine is so large that it cannot be nullified by increasing α, and x_h will be set equal to ψ.

Proof: See appendix.

Proposition 4 shows the limitations of a fixed fine on the manager in reducing harm just as Proposition 2 shows the limitations of a fixed fine on the corporation in reducing the harm it causes. First, for low levels of the fine, the manager will choose the same level of the harmful activity that is chosen without the fine and this requires no change in his compensation. For intermediate levels of the fine, the manager's compensation can be adjusted so that he continues to choose the level of the harmful activity he would choose in the absence of a fine. Only when the fine is very significant does it have an effect.

Based again on the parameters from our Numerical Example, Figure 4 shows the α that must be set for the manager to choose $x_h = x_h^*$ for a range of levels of F_m from 0.00 to 0.02. It is assumed that $\psi = 0$, which means that the manager is fined for any value of $x_h > 0$. Recall that the manager's utility when there is no fine is 0.01 (in monetary terms), so at the upper range of what is shown the fine is quite substantial. As the figure shows, for low levels of the fine, α can be set to the level it would be without a fine, 0.4%. Once the fine increases to the level of 0.004, α needs to be adjusted upward to motivate the manager to set $x_h = x_h^*$ and incur the fine. Note that in this example even at the highest level shown, the adjustment in α is not that significant.

fine on the corporation will be no greater than just leveling the fine on the corporation.

One can show that relative to shareholder profits when α is left unchanged at 0.4%, shareholder profits can be almost 35% higher when they adjust α even when accounting for shareholders' smaller share of the profits due to increasing the equity compensation of the manager. It should be noted, however, that α must be increased much more when $x_h^* - \psi$ is small. That is because the increase in profits is not that large when x_h is increased from ψ to x_h^* and the fine is incurred. Thus, a larger share of that increase must be given to the manager to create incentives to choose x_h^* and incur the fine.

We now examine what happens when a variable fine equal to $f_m\sqrt{x_h}$ is levied on the manager. As stated in Proposition 5, as long as the fine is not too large, the corporation can adjust the manager's compensation to partially offset the effects of the fine so that the manager chooses a higher level of the profitable but harmful activity.

Proposition 5. *There exists an $\bar{f}_m$ such that when $f_m < \bar{f}_m$, the corporation will adjust the manager's compensation to provide incentives to the manager to increase the level of the harmful activity. When $f_m \geq \bar{f}_m$, then the corporation will adjust the manager's compensation so that $x_h^* = 0$, and harm is fully deterred.*

Proof: See appendix.

The intuition behind Proposition 5 is that when a fine of $f_m\sqrt{x_h}$ is imposed on the manager, the harmful activity becomes less rewarding for the manager. The manager does cut back on the harmful activity, but at the level of α that the corporation establishes when there is no fine, the manager does not have the incentive to increase x_b to the point where the chosen activity levels sum to one. The manager needs a bigger fraction of the profits to motivate the profit-maximizing choices. The corporation therefore increases α which in turn increases the attractiveness of the harmful activity and therefore increases the level x_h that is chosen. Figure 5, which is based on the parameters given in our Numerical Example, illustrates this.

By imposing a fine directly on the manager, a much smaller fine can be used to deter corporate misconduct. When a fine is imposed on the corporation, only a fraction α flows through to affect the manager, whereas a managerial fine is imposed entirely on the manager and can potentially have a much stronger effect on his incentives. For a high enough variable fine, misconduct will be deterred.

However, in our analysis of fines levied against the manager so far we have assumed that the corporation can only respond by changing the equity component of the manager's compensation. It would often be more efficient and less costly for the corporation to respond by committing to pay the full amount of any fines imposed on the manager, which would motivate him to make the same choices that would be made without any fines. Indeed, managerial contracts routinely include substantial insurance and indemnification (Schwab and Thomas, 2006).[26] Whereas it is not generally possible to insure managers against criminal charges, firms can—and regularly do—insure and indemnify managers against civil liability, and in some cases help them defend against criminal liability directly or indirectly. That is, indemnification effectively turns a managerial fine into a very small corporate fine. This observation illustrates again our general point: if external law enforcement is imperfect, internal governance may be designed to exploit those imperfections and often in ways that are not all that costly.[27]

[26]For example Alphabet has "entered into an indemnification agreement with each of our directors and executive officers. The indemnification agreements, our certificate of incorporation, and bylaws require us to indemnify our directors and executive officers to the fullest extent permitted by Delaware law."

[27]It is important to note, of course, that insurance and indemnification also serve the important role of protecting

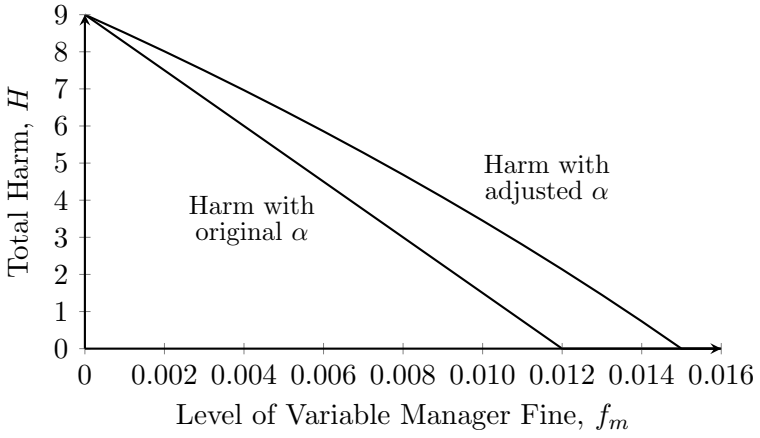


Figure 5: **Harm as a Function of a Variable Manager Fine and Equity Adjustment.** The plot shows the level of harm for a given level of a fine for both the original level of $\alpha = 0.004$ and for the optimal level of α for the parametrization $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $h = 15$. This shows that the increase in managerial compensation leads to greater levels of harm.

The US Department of Justice has emphasized an increased focus on individual-level liability. In other jurisdictions (e.g., Germany), enforcement actions tend to focus on individuals and sanctions imposed on the corporation are limited. From a narrow reading of the above propositions, the focus seems appropriate. That is, it seems that by imposing liability directly on the manager, the government can achieve the same amount of deterrence for a much lower fine. However, this logic ignores the possibility of indemnification and insurance. Indeed the use of indemnification as well as insurance are pervasive in managerial (and director) compensation, and such practices can further weaken external enforcement and deterrence.

To see this point more clearly, suppose that the government is considering imposing a corporate fine, f_c, or a managerial fine, f_m, where $f_m < f_c$. Absent insurance and indemnification, which is the better option will depend on the levels of the fines along with a wide variety of other factors including managerial compensation, the disutility of effort, and the profitability of both the beneficial and harmful actions. For a given set of parameter values, either could be better from a social welfare perspective. However, it is also true that for any corporate-level fine, f_c, there exists some manager-level fine $f_m \leq f_c$ that achieves at least as good of an outcome. If it is more feasible to impose the smaller fine on the manager than the larger fine on the corporation, then the manager-level fine may be more appropriate, consistent with the Department of Justice's policies.

However, when indemnification is possible, imposing liability $f_m < f_c$ on the manager is worse than imposing liability f_c on the corporation. This conclusion follows straightforwardly, because indemnification turns a managerial-level fine into a corporate level fine. This illustrates the dangers of ignoring internal governance responses to external governance changes. If corporations are able to insure and indemnify their managers, then the increased focus on managerial fines

managers and directors from frivolous lawsuits. Under Delaware Corporate Law, coverage must exclude any "deliberate criminal or deliberate fraudulent act" or "knowing violation of law," DGCL 145(g). However, this exclusion only applies when the loss is established by a "final, non-appealable adjudication in the underlying proceeding in respect of the claim," which rarely occurs. Mullin and Snyder (2010) formally model some roles that indemnification might play. They show that if the government can mistakenly prosecute innocent managers, indemnification can be beneficial if it transfers that risk from risk-averse agents to risk-neutral principals.

may in fact reduce deterrence and exacerbate corporate misconduct. In principle, indemnification completely nullifies the effect of managerial fines if it is complete and certain. However, it can still be effective in offsetting some of the effects of fines when it is achieved in tacit and less direct ways, such as implicit promises or understandings to pay executives more after the imposition of individual liability.

Of course, beyond imposing fines the state can threaten criminal prosecution against the manager, which in serious cases may result in jail time. While the corporation might increase managerial pay to partially compensate for criminal penalties and make the manager willing to take this risk, the deterrence effects could be strong. However, in considering the potential for criminal prosecution to deter corporate misconduct, it should be noted that in most cases the bar for conviction is high. Governments have often been reluctant to prosecute executives for corporate misconduct and unsuccessful in achieving convictions and jail sentences when they do(Rakoff, 2021; Buell, 2024).[28]

The difficulty in creating deterrence through criminal prosecution is partly due to the fact that corporations can and may well act to weaken these deterrent effects if they can reduce the likelihood of the manager being charged. When setting internal governance policies, it is in the corporation's interest, to the extent it finds the misconduct profitable, to shield the manager from any legal consequences in order to realign the interests of the manager with the corporation. An example discussed in subsectionis Volkswagen, which paid to settle criminal charges against two executives, and the role of insurance and indemnification to shield against civil liabilities. As discussed in Eisinger (2017), it is common for US corporations to cover high-cost legal defense for executives and for executives to coordinate their defense strategies with those of the corporation.

Throughout our analysis we have assumed that the manager experiences the same disutility of effort in pursuing either the beneficial or the harmful action at a given level. It could be objected that some managers may have moral qualms about the harmful action and experience more disutility in pursuing it. If the disutility associated with the harmful action is increased by these moral qualms, the manager will choose to pursue it at a lower level for any given compensation contract. However, just as the corporation can increase the manager's equity stake to induce him to engage in the harmful activity when the manager faces fines, the corporation can potentially adjust compensation to overcome some of the manager's moral qualms. While hiring managers with moral convictions can help improve social welfare, a corporation completely focused on profits would try to avoid hiring such a manager.

3.4 Tying Penalties to Harm

Up to this point we have assumed that h, the *rate* of the harm caused by the harmful activity per unit of $\sqrt{x_h}$, is fixed; the corporation only has control over x_h, the *level* of the harmful but profitable activity. We now consider situations in which the corporation can take actions that reduce the rate of harm h. When the size of the fines imposed is based on the level of profits the corporation gains from the harmful activity, as we have generally assumed above, the corporation will have no incentive to reduce the rate of harm. For this reason, one might think that it would be better to relate fines to the harm done rather than to profits gained, since this would give incentives to the corporation to reduce the rate of harm. After all, as has been argued by, for example, Becker (1968), making fines equal the *total* harm caused by the activity can

[28]The recent cases of Elizabeth Holmes of Theranos or Sam Bankman-Fried of FTX are the exception that proves the rule. In both cases, the corporations were closely held startups that went into bankruptcy, and their outspoken CEOs provided unusually strong evidence that enabled prosecutors to achieve convictions.

lead to efficient and welfare maximizing outcomes. As we now show, however, when fines are insufficiently high, basing fines on the level of harm can perversely *reduce* social welfare even though the corporation is motivated to take actions that reduce the *rate* of harm.

To explore the issues in a specific context, we assume that the corporation is subject to a variable fine. As shown in Proposition 3, in this case the manager's compensation can be set so that his decisions maximize the corporation's profits. We will therefore refer throughout this subsection to choices that maximize the corporation's profits knowing that the corporation will adjust compensation to account for the fine so as to ensure that the manager makes these choices.

We modify the model in the following manner. Let the baseline rate of harm be h_0. We assume that there some mitigating actions, potentially costly, that the corporation can take to reduce the level of harm and that the rate of harm when harm mitigation action k is taken is h_k with $h_0 > h_1 > h_2 > \ldots > h_K$.

We begin by assuming that the actions that reduce the rate of harm are costless. If fines are only imposed at the corporate level, the corporation's objective function is:

$$\max_{x_b, x_h, k} \left(\pi_b \sqrt{x_b} + \left(\pi_h - f_c \times h_k \right) \sqrt{x_h} \right) \quad \text{s.t.} \quad x_b + x_h \leq 1, \quad x_b, x_h \geq 0 \tag{6}$$

It is quite clear that when there is no cost to reducing the rate of harm, the corporation will choose k to minimize h_k, since this minimizes the fine imposed for any given level of x_h. Will this increase social welfare?

Consider the Numerical Example we have used throughout with $f_c = 0.15$. The baseline rate of harm is $h_0 = 15.0$, but the corporation can take mitigating actions to reduce harm by 10% or 20%. Thus, the possible rates of harm are $\{h_0, h_1, h_2\} = \{15.0, 13.5, 12.0\}$. Clearly, if x_h is held fixed, then reducing the rate of harm from 15.0 is unambiguously beneficial for society. However, the corporation's choice of mitigation also affects the corporation's optimal choice of x_h. The upper panel of Figure 6 shows the corporation's profits as a function of x_h for the three levels of h_k. The vertical lines indicate the profit-maximizing levels of x_h.

Observe that since mitigation leads to lower fines for any given level of x_h, the profit-maximizing level of x_h increases with mitigation. Of course, at the same time, mitigation decreases the total harm caused at any *given* level of x_h. The lower panel of Figure 6 shows that given the corporation's profit maximizing incentives, the total harm, $h_k \sqrt{x_h}$, will increase when mitigation is possible. The dashed lines show the profit-maximizing levels of x_h for each level of h_k, and it is easily seen that the corporation's choice of x_h for each h_k means that total harm will increase as the corporation reduces the rate of harm. In this case, tying fines to harm increases the total harm caused relative to the base rate of $h_0 = 15.0$.

To understand the intuition behind this result it is useful to consider the intensive and extensive margins that determine the total harm. The rate of harm, h_k, defines the intensive margin, while the level of the harmful activity, x_h, defines the extensive margin. Making the fine depend on total harm, which in our model is $H = h_k \sqrt{x_h}$, encourages the corporation to choose a lower rate of harm, which moves the intensive margin in a socially beneficial direction. However, reducing the rate of harm means that the profitability of the activity, $\pi_h - f_c \times h_k$, increases, which leads the corporation to increase the level of the harmful activity x_h along the extensive margin. The total harm depends on how much the increase on the extensive margin offsets the intensive margin reduction. It is quite possible for total harm to increase.

We have assumed that mitigation is not costly. However, since profits increase when the rate of harm is reduced, it is clear that the corporation will be quite willing to take costly mitigating actions to reduce h to 13.5 or 12.0 as long as these actions are not too costly.

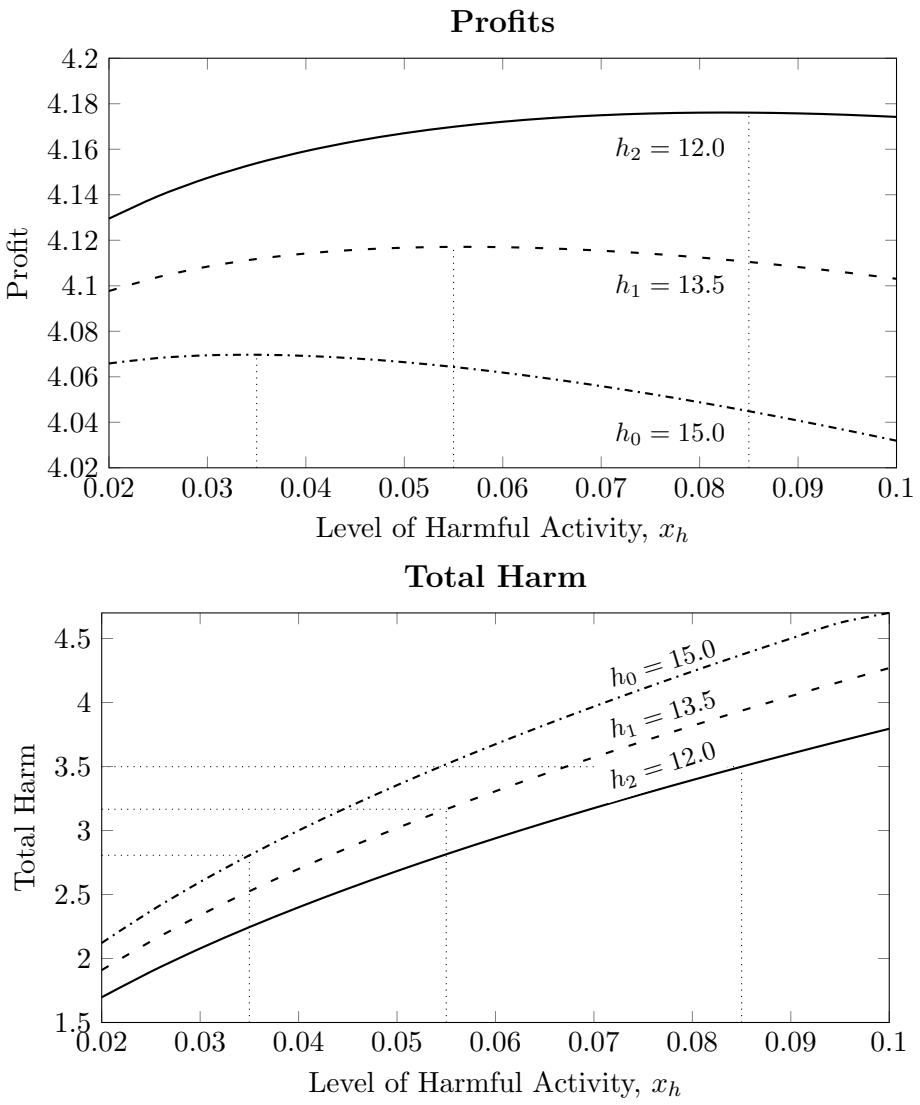


Figure 6: **Profits and Harm as a Function of a Harmful Activity with Mitigation**. For the parameters $\pi_b = 4$, $\pi_h = 3$, $\beta = 0.01$, $b = 6$, and $f_c = .15$, the upper plot shows corporate profits for $h_0 = 15.0$, $h_1 = 13.5$ and $h_2 = 12.0$ as a function of the harmful activity,. The lower plot shows the total level of harm as a function of the harmful activities for the three rates of harm. The dashed vertical lines correspond to the profit maximizing levels of x_h.

In an ideal world, f_c would be high enough to fully deter the harmful activity so that the corporation would set $x_h = 0$. However, when enforcement is weak and fines fall below these levels, tying fines directly to harm can produce the unintended consequence of actually increasing the total harm. This perverse outcome is not a mere knife-edge curiosity. Instead, whenever monetary sanctions are set too low, it is quite possible that the effect of well-intended policies will increase harm and decrease social welfare, underscoring again that the core problem is inadequate penalties.

4 Cooperation and Discounted Fines

Our analysis has so far considered a single period in which the total amount of harm was completely determined by the rate of harm h and the the level of harmful activity, x_h. We also assumed that any misconduct was detected with certainty. This setting was sufficient for exploring how the deterrent effect of fines can be reduced by actions taken by the corporation, particularly ones that change the incentives of the decision-making manager.

In reality, corporate misconduct is rarely an event that occurs just once, but instead typically involves many decisions made over a long period of time. This means that a major determinant of harm is how long the harm remains undetected and is allowed to persist. Harms can remain hidden from both law enforcement and from victims for many months or years. Detection of misconduct often relies on chance events such as plane crashes, whistle-blowers speaking out, or media investigation (see, e.g., Dyck, Morse and Zingales (2010) on fraud). Some misconduct is only uncovered when a corporation defaults on debt obligations or files for bankruptcy (for example, as in the cases of Enron, Madoff, Theranos and FTX) .

Since detection and investigations are difficult and costly, authorities often encourage corporations to put in place compliance programs that would contribute to earlier or more detection. They also encourage firms to self-report misconduct to authorities. This encouragement takes the form of offering corporations discounted fines and reductions in other penalties and sanctions. To examine this possibility we develop in this section a multi-period model where in each period the probability of the harmful activity being detected is less than 1. We consider two ways that have been proposed or are already used in practice to increase the probability of detection and reduce the duration of harm.

The first approach that we examine involves encouraging corporations to voluntarily adopt a compliance program in return for reductions in fines when misconduct is detected. We assume adopting such a compliance program brings about changes in internal monitoring and increases in transparency that ultimately mean misconduct is more likely to be detected by the authorities. The second approach is an extreme form of this idea, where the corporation self-reports the misconduct, again motivated by a reduction in fines. If the reduction in the fine is significant enough to make self-reporting desirable for the corporation, the probability of detection is equal to 1 when self-reporting takes place. As throughout our analysis, our focus is on cases where the corporations and its shareholders benefit from the misconduct while society is harmed.

In the next two subsections we will analyze compliance programs and self-reporting in a multi-period setting. We restrict our attention throughout our analysis to the case where the corporation is subject to the variable fine f_c, and we modify the model in the following manner:

1. **Timing**. There is an infinite number of periods indexed $k = 1, 2, \ldots$. Profits and the monetary value of harms occurring in period k are discounted to the present by δ^k with $\delta < 1$.

2. **Probability of Detection.** At the start of each period there is an exogenously given probability $\bar{p}$ that the corporation's misconduct will be detected if it has engaged in any misconduct and has done nothing to increase the probability of detection. (The specific steps the corporation can take to increase the probability of detection are discussed in Sections 4.1 and 4.2 respectively.)

3. **Fines**. If misconduct is detected at the start of period k, the corporation pays a fine based on current value of its total accumulated profits up through period $k-1$:[29]

$$F_k = f_c\sqrt{x_h^\star} \times \frac{1-\delta^{k-1}}{\delta^{k-1}(1-\delta)}. \tag{7}$$

4.1 Voluntary Compliance Programs

In this subsection we extend the multi-period model to take into account the possibility that authorities offer discounts in fines to corporations that choose to implement compliance programs. Our model captures the idea that (1) some aspects of compliance programs are voluntary, (2) compliance programs will genuinely increase the probability of detection and a fine being incurred, and (3) implementing compliance programs can lead to discounts in fines. When we refer to voluntary compliance programs, the interpretation will include the anticipated reduction in fines.

We have three main results in this section. First, because voluntary compliance programs reduce the rate of the fine, they will generally increase x_h, the levels of the harmful activity in any period. Second, the effects of voluntary compliance programs on the expected total harm caused by the misconduct are generally ambiguous — the programs may either increase or decrease total harm. Third, when the opportunity to engage in misconduct is only available in one period (rather than every period), offering discounts for voluntary compliance programs will never reduce the harm caused and can actually increase the harm.

Specifically, we add the following two assumptions to the multi-period model above to model compliance programs that improve detection of misconduct in exchange for discounted fines upon detection:

1. **Choice of Compliance Program**. The corporation can voluntarily implement a compliance program that *increases* the detection probability from $\bar{p}$ to some level $p \in (\bar{p}, 1)$.

2. **Discounted Fine**. If the corporation implements $p > \bar{p}$, the government reduces the fine from F_k to $\big(1 - D(p)\big) \times F_k$, where $D(p)$ equals 0 when $p = \bar{p}$ and is increasing over the range $p > \bar{p}$.

Note that we assume that the size of the discount increases in the effectiveness of the compliance program that the firm implements in facilitating detection of misconduct. Implementing a voluntary compliance program presents the corporation with a tradeoff. If the corporation implements a program, it will increase the probability of detection, which *decreases* the expected number of periods the corporation can earn profits from misconduct. However, the discounts that motivate that program reduce the effective fine imposed on the corporation for any given level of misconduct, which *increases* the profitability of misconduct in any period wrongdoing has not been detected. This effect leads the corporation to increase the level of misconduct x_h

[29] We are assuming that when misconduct is detected, the fine imposed at that time is based on the *current* value at that time of all profits that were received by the corporation due to the misconduct. This means that the fine on profits received j periods before detection is grossed up by a factor of $1/\delta^j$.

above what it would be with no discount. There is thus an "extensive margin" and an "intensive margin" that work in opposite directions. Whether the program increases social welfare depend on which dominates. Critically, the program is *voluntary*, so the corporation will only implement the program if it expects to profit.

We first consider the conditions under which a corporation will voluntarily adopt a program. Define Π^M to be the maximum per-period profits the corporation can earn if it can engage in misconduct and the fine discount is $D(p)$:

$$\Pi^M = \pi_b \sqrt{x_b^\star\left(D(p)\right)} + \Big(\pi_h - (1 - D(p))f_c\Big)\sqrt{x_h^\star\left(D(p)\right)}, \tag{8}$$

where the optimal activity levels are functions of $D(p)$ and are equal to:

$$x_b^*\big(D(p)\big) = \frac{\pi_b^2}{\pi_b^2 + \big(\pi_h - \big(1 - D(p)\big)f_c\big)^2}; \quad x_h^*\big(D(p)\big) = \frac{\big(\pi_h - \big(1 - D(p)\big)f_c\big)^2}{\pi_b^2 + \big(\pi_h - \big(1 - D(p)\big)f_c\big)^2}. \tag{9}$$

Note that in the infinite multi-period setting we are analyzing, the misconduct will be detected with probability 1 at some point as long as $\bar{p} > 0$. Thus, the corporation assumes that fines will be paid with certainty.

Since by assumption, $\pi_h > f_c$, we know that $\pi_h > \big(1 - D(p)\big)f_c$ and $x_h^\star\big(D(p)\big) > 0$. It follows that for any $D(p) \in [0, 1]$, $\Pi^M > \pi_b$, where π_b is the profit the corporation will earn if it does not engage in misconduct. So the optimal strategy for the corporation will be to engage in misconduct until it is detected, regardless of whether or not the corporation introduces a voluntary discount program.

Begin by noting that for since $D(p) > 0$ whenever the firm adopts a compliance program, it follows that for all $p > \bar{p}$:

$$x_h^*(D(\bar{p})) = \frac{(\pi_h - f_c)^2}{\pi_b^2 + (\pi_h - f_c)^2} < \frac{\big(\pi_h - \big(1 - D(p)\big)f_c\big)^2}{\pi_b^2 + \big(\pi_h - \big(1 - D(p)\big)f_c\big)^2} \equiv x_h^*(D(p)). \tag{10}$$

That is, introducing discounts for voluntary compliance programs *always* increases the per-period choice of the harmful activity. The only way introducing discounted fines in exchange for compliance programs can be socially beneficial is if the increased level of harmful activity is more than offset by the increased probability of detection.

The discounted value of the corporation's expected profits from engaging in misconduct is:

$$\sum_{k=1}^{\infty} \delta^k (1-p)^k \Pi^M + \sum_{k=1}^{\infty} \delta^k p (1-p)^{k-1} \left(\Pi^M + \frac{\delta \pi_b}{1-\delta}\right). \tag{11}$$

The first sum is the discounted value of all the profits earned in periods when the misconduct is not detected. The second sum accounts for periods when the misconduct is detected and the corporation gets both its final payoff from misconduct, Π^M, and the continuation value of the future stream on profits with no misconduct, $\delta\pi_b/(1-\delta)$. The expected value in (11) can be simplified to:

$$\frac{\delta}{1 - \delta(1-p)} \times \Pi^M + \frac{\delta p}{1 - \delta(1-p)} \times \frac{\delta \pi_b}{1-\delta}. \tag{12}$$

The present value of social harm based on the expectation of all misconduct is

$$\text{Present Value of Expected Harm} = \frac{\delta h \sqrt{x_h^\star(D(p))}}{1-\delta(1-p)}. \tag{13}$$

It can now be seen that the option of implementing a voluntary compliance program has an ambiguous effect on the expected harm. The increase the probability of detection, which makes earlier detection more likely, will reduce the denominator and, all else equal, reduce social harm. However, the discount in the level of the fines imposed when p is increased will increase the profitability of the harmful action, thereby increasing $x_h^\star$ and thus the numerator. If the effect of the increased probability of detection dominates, then voluntary compliance programs will reduce social harm. If instead the effect of the increase in the harmful action dominates, then voluntary compliance programs will exacerbate social harm. Regardless of the effect on the overall social harm, *voluntary* compliance programs lead to an increase in corporation's expected net profits from misconduct. If this were not true, the corporation would not voluntarily implement the program.

Equation (13) leads directly to the following proposition.

Proposition 6. *Offering a discount for the implementation of voluntary compliance programs has an ambiguous effect on total social harm.*

To illustrate the consequences of this ambiguity, consider the parameters in our Numerical Example used throughout ($h = 15$, $\pi_b = 4$ and $\pi_b = 3$). Assume that $\delta = 0.90$ and that the base-rate probability of detection is $\bar{p} = 0.4$. The corporation can take actions to increase p, and for every 0.05 in increase in p the fine is discounted by an additional 10%, but the maximum discount is 40%.

Table 2 shows the effect of various voluntary compliance programs when the fine is $f_c = 2$. The first row of the table shows the outcome if the corporation does not adopt a compliance program and the probability of detection remains at 40%. Each subsequent row shows a different level of compliance program that translates to an increased detection probability. As p increases, the expected duration of harm goes down, but the level of the harmful activity increases. The corporation chooses p to maximize its profits. In this example, $p^* = 0.6$, for which the corporation receives the maximum 40% discount on its fines. Even though the compliance program cuts the expected duration of profitable misconduct by a third, it leads the corporation to increase the harmful activity by nearly three fold, which more than doubles the corporation's profits from misconduct.

Table 2: Fine = 2.0

p	Expected Duration of Harm	Discount on Fine	x_h^*	Profits Above No Misconduct	Expected Harm	Expected Harm If Mandatory
0.40	2.500	0%	0.059	0.241	7.118	–
0.45	2.222	10%	0.083	0.314	7.682	6.484
0.50	2.000	20%	0.109	0.389	8.109	5.953
0.55	1.818	30%	0.138	0.466	8.427	5.503
0.60	**1.667**	**40%**	**0.168**	**0.543**	**8.656**	**5.116**
0.65	1.538	40%	0.168	0.508	8.087	4.780

Importantly, the compliance program increases the level of misconduct, corporate profits, and social harm because it is *voluntary*. That is, the corporation chooses the detection probability to

maximize profits given the enforcement policy. If instead the corporation were *required* to implement a compliance program that resulted in a 60% probability of detection (without benefiting from a discount), the expected harm would be much lower, as can be seen in the last column of the table. Critically, therefore, it is the voluntary nature of the compliance program, which uses the discounts to provide incentives for the corporation, that can make relying on corporations to increase the probability of detection counterproductive.

Some enforcement policies in the real world capture the idea of a mandatory compliance programs. For example, power plants have mandatory hourly reporting requirements for key pollutants, and rather than being rewarded for having a functioning reporting regime, they are penalized for not having one in place. However, most corporations are not subject to such specific compliance and reporting requirements. Our results show that simply reducing fines in order to motivate corporations to detect and disclose misconduct with higher probability can makes matters worse if the misconduct benefits shareholders

Voluntary compliance programs do not always increase social harm. In Table 3 we assume that the baseline fine level is lower, specifically that $f_c = 1$. In this case, the corporation's profit-maximizing choice is again to set $p = 0.6$, entitling it to the 40% discount. The corporation will once again respond to the fact that the fine is lower by increasing the level of the harmful activity. Because the baseline fine is so much lower in this example, however, the gain from the discount is much lower than when the baseline fine was $f_c = 2$, so the harmful activity level increases by much less. In this case, therefore, the benefits from reducing the duration of the misconduct outweigh the costs of increasing the harmful activity level, so that the expected harm is lower. Again, a mandatory compliance program with no discounts would be much better—and increasing the fine would be better still, as seen in Table 2

Table 3: Fine = 1.0

p	Expected Duration of Harm	Discount on Fine	x_h^*	Profits Above No Misconduct	Expected Harm	Expected Harm If Mandatory
0.40	2.500	0%	0.200	0.924	13.125	–
0.45	2.222	10%	0.216	0.923	12.426	11.955
0.50	2.000	20%	0.232	0.925	11.829	10.977
0.55	1.818	30%	0.248	0.929	11.310	10.147
0.60	**1.667**	**40%**	**0.265**	**0.935**	**10.853**	**9.433**
0.65	1.538	40%	0.265	0.873	10.140	8.814

It is also possible that the discounts offered are such that, relative to the exiting fine, the corporation will not voluntarily adopt a compliance program. Table 4 shows the same parameters with a fine of $f_c = 0.5$. In this case, the corporation will leave $p = \bar{p} = 0.4$, because the benefit of the discount on the fine is too small relative to the corporation's costs from the lower expected duration of misconduct.

This analysis also calls into question the "best" policy choices. Offering discounts for voluntary compliance programs has no effect when the fines are already very low ($f_c = 0.5$), and has a small beneficial effects when fines are somewhat low ($f_c = 1$). However, when fines are relatively high ($f_c = 2.0$), offering discounts for voluntary compliance programs significantly increases social harm. That is, in the context of this example, offering discounts to encourage compliance programs is only justifiable if fines are set much too low. If that is the case, it would be better to simply increase the fines.

Table 4: Fine = 0.5

p	Expected Duration of Harm	Discount on Fine	x_h^*	Profits Above No Misconduct	Expected Harm	Expect Harm If Mandatory
0.40	**2.500**	**0%**	**0.281**	**1.403**	**15.554**	–
0.45	2.222	10%	0.289	1.325	14.370	14.168
0.50	2.000	20%	0.297	1.261	13.377	13.009
0.55	1.818	30%	0.305	1.207	12.531	12.025
0.60	1.667	40%	0.313	1.162	11.801	11.180
0.65	1.538	40%	0.313	1.085	11.026	10.445

Ultimately, because voluntary compliance programs affect social harm on two margins, it is hard to make any accurate predictions about whether they can be expected to increase or decrease social harm.

Whereas in the above discussion we examined a multi-period model of misconduct, in some cases a one-period model is sufficient because the misconduct is a "one-off" opportunity available to the corporation in a single period. Then we have the following stark result:

Proposition 7. *If the opportunity for misconduct is only available in a single period, compliance programs will either have no effect on the level of harm or will increase the level of harm.*

Proof: See appendix.

The intuition behind this result is straightforward. In a situation where the harm can only occur in one period, any increase in the probability of detection only increases the chance the corporation will pay the fine — it does nothing to reduce the *duration* of the harmful activity. The expected fine the corporation faces is $p \times (1 - D(p)) \times f_c$. The corporation will choose to engage in the voluntary compliance program only if there is a $p > \bar{p}$ such $p \times (1 - D(p)) \times f_c < \bar{p} \times f_c$. If no such p exists, the corporation will not voluntarily implement the compliance program and the possibility of implementing the program will have no effect. If there is a $p > \bar{p}$ that reduces the expected fine, the corporation will gladly implement the voluntary compliance program and increase x_h since doing so will make the misconduct more profitable net of the discounted fine. In this case the voluntary compliance program has the effected of *increasing* social harm.

Our discussion of compliance programs has focused on the corporation rather than manager. However, our results from Section 3 show that the corporation can, and will, adjust the manager's compensation to align the manager's decisions with the corporation's profit calculations. If the corporation finds that "good" compliance (i.e., a higher p accompanied by a large discount) increases profits, it can generally ensure that the manager's profit-based pay provides the incentives needed to choose the profit maximizing strategy. As with any external enforcement mechanism, the final outcome is determined by the corporation's private trade-off between the cost of misconduct delivered through fines and sanctions and the profitability of harmful activity.

Finally, in our analysis of compliance programs in this section, we have assumed that they are effective in increasing the probability of detection. However, in some cases the corporation may create the *appearance* of an effective compliance program without meaningfully increasing the probability of detection as we assume in our model. In this case the compliance program will almost certainly increase harm since the discount will increase x_h which increases the numerator

of (equation with harm) but will have little if any effect on increasing the denominator, since it does not materially decrease the expected duration of the harm.

4.2 Self-Reporting Programs

We now turn to enforcement policies that encourage corporations to self-report past misconduct by promising a reduction in fines. In the model of self-reporting we discuss below, we assume that in any period in which the corporation does not report its misconduct the probability of detection is fixed at an exogenously given level $\bar{p}$. If the corporation chooses to self-report in a given period, the probability of detection in that period jumps to 100%. In the voluntary compliance program we analyzed above the corporation makes a one time decision at the beginning whether or not to participate in the program. With self-reporting, the corporation must decide each period whether to admit its misconduct and pay a reduced fine or stay silent and gamble that it can get away with the misconduct for additional periods without being detected and having to pay the full fine.

Our results in this section mirror those of Section 4.1. First, discounts for self-reporting increase the corporation's incentives to engage in the harmful activity. Second, the effect of offering discounts for self-reporting on the total harm caused by misconduct is generally ambiguous—self-reporting programs may either increase or decrease the total amount of harm caused. Third, when misconduct is chosen only once (rather than every period), offering discounts for self-reporting will sometimes increase but never decrease harm.

We assume that at the beginning of each period the corporation can self-report misconduct that occurred in prior periods and we assume that all misconduct is reported. The corporation then will pay a *discounted* fine based on the current value of the profits it earned in prior periods due to misconduct. If the corporation does not self-report at the beginning of the period, then with probability $\bar{p}$, its prior misconduct will be detected and it will be required to pay the full amount of the fine. After the misconduct is either self-reported or by chance detected, the corporation cannot engage in any misconduct in that period or in any future period. If it does not self-report and misconduct is not detected this period, the corporation can continue to engage in and earn profits from the harmful activity until some future period when the misconduct is either self-reported or by chance detected. Formally, we add the following assumptions to the multi-period model:

1. **Self-Reporting**. At the beginning of each period k, the corporation can choose whether or not to self-report. If it self-reports, it reveals all misconduct from prior periods $1, \ldots, k-1$ and must set $x_h = 0$ for period k and all subsequent periods. It will be required to pay a fine in period k equal to $(1 - D)F_k$.

2. **Keeping silent.** If the corporation does not self-report in period k, then there is probability $\bar{p}$ that all its misconduct is revealed in another way. If misconduct is revealed, it again must set $x_h = 0$ for period k and all subsequent periods. In this case, however, it must pay the full fine of F_k. If the firm is "lucky" and its misconduct is not revealed, something that occurs with probability $1 - \bar{p}$, it can engage in misconduct in period k, setting $x_h = x_h^\star$ and earning higher profits than when $x_h = 0$.

To understand when the corporation finds it optimal to self-report, note first that the value of self-reporting in period k is:

$$\text{Report}(k) = \frac{\pi_b}{1-\delta} - (1 - D) \times F_k. \tag{14}$$

Upon self-reporting, the corporation receives the period k present value of profits that are earned without engaging in misconduct from period k onward (the first term) but gets the benefit of paying the *discounted* accumulated fine (the second term).

The value of staying silent in period k and self-reporting in the subsequent period is:

$$\text{Silent}(k) = \bar{p} \times \left(\frac{\pi_b}{1-\delta} - F_k \right) + (1-\bar{p}) \Big(\Pi(x_b^\star, x_h^\star) + \delta \times \text{Report}(k+1) \Big). \tag{15}$$

If the corporation stays silent and gambles it will not be detected, then with probability $\bar{p}$ its misconduct is detected and its loses that gamble. It must pay the full fine and gets the present value of profits attained without engaging in misconduct thereafter (the first term). However, with probability $1-\bar{p}$ it wins the gamble and evades detection. When undetected it can generate higher profits in that period equal to $\Pi(x_b^\star, x_h^\star)$, and then, assuming it self-reports the next period, gets the the value of self-reporting in period $k+1$ discounted to period k.

The corporation will self-report in the first period k that $\text{Report}(k) > \text{Silent}(k)$. Note that as long as both $\bar{p}$ and D are greater than zero, this condition will be satisfied for some k since F_k grows without bound. Because the corporation will always self-report at some point, the expected fine from misconduct is smaller for a given choice of x_h under a self-reporting regime. It follows that the corporation will always choose a higher level of the harmful action when discounts in fines for self-reporting are offered than when no discounts are offered. That is, because of the discount, self-reporting regimes always increase the corporation's level of harmful activity. Just as is true with the voluntary compliance programs, self-reporting operates on two margins. It increases the corporation's per-period level of the harmful activity, but decreases the expected duration of the harmful activity, which is socially beneficial. Once again the effects are ambiguous

Proposition 8. *Offering a discount for self-reporting has an ambiguous effect on total social harm.*

To see how self-reporting can either reduce or increase social harm, consider the Numerical Example we have used above: $h = 15$, $\pi_b = 4$, $\pi_h = 3$, $\delta = 0.90$, and $\bar{p} = 0.4$. Consider first the case where fine rate f_c is equal to 2.0 and the corporation receives a 50% discount in the accumulated fine if it self-reports. Panel (A) of Figure 7 shows the corporation's expected profits above what it would have realized if it never engages in misconduct (i.e., set $x_h = 0$ for all k) as a function of self-reporting in period k. In other words, these are the extra profits it obtains by engaging in misconduct and then self-reporting at a given period. The horizontal dashed line represents the corporation's expected profits if no discount is offered for self-reporting, in which case the firm will engage in misconduct until it is caught.

The first period the corporation can self-report is period 2, when it reports its misconduct in period 1. Relative to never self-reporting, this increases the corporation's profits. However, from its perspective it pays for the corporation to gamble in period 2 and stay silent since a strategy of self-reporting in period 3 yields higher expected profits. In fact, this strategy yields the highest expected profits — staying silent and self-reporting in period 4 (and any subsequent period) is dominated by self-reporting in period 3, which is what the corporation will do.

Panel (B) of Figure 7 shows x_h^*, the amount of harm the firm would do each period for the various self-reporting strategies it might follow. Note that x_h^* declines as the self-reporting period is put further in the future since staying silent longer increases the expected fine rate since the firm is gambling more on not being discovered. The probability that it will be detected before it self-reports and be required to pay the full fine increases with each additional period it gambles on its misconduct not being discovered. The dashed line is the per-period level of misconduct

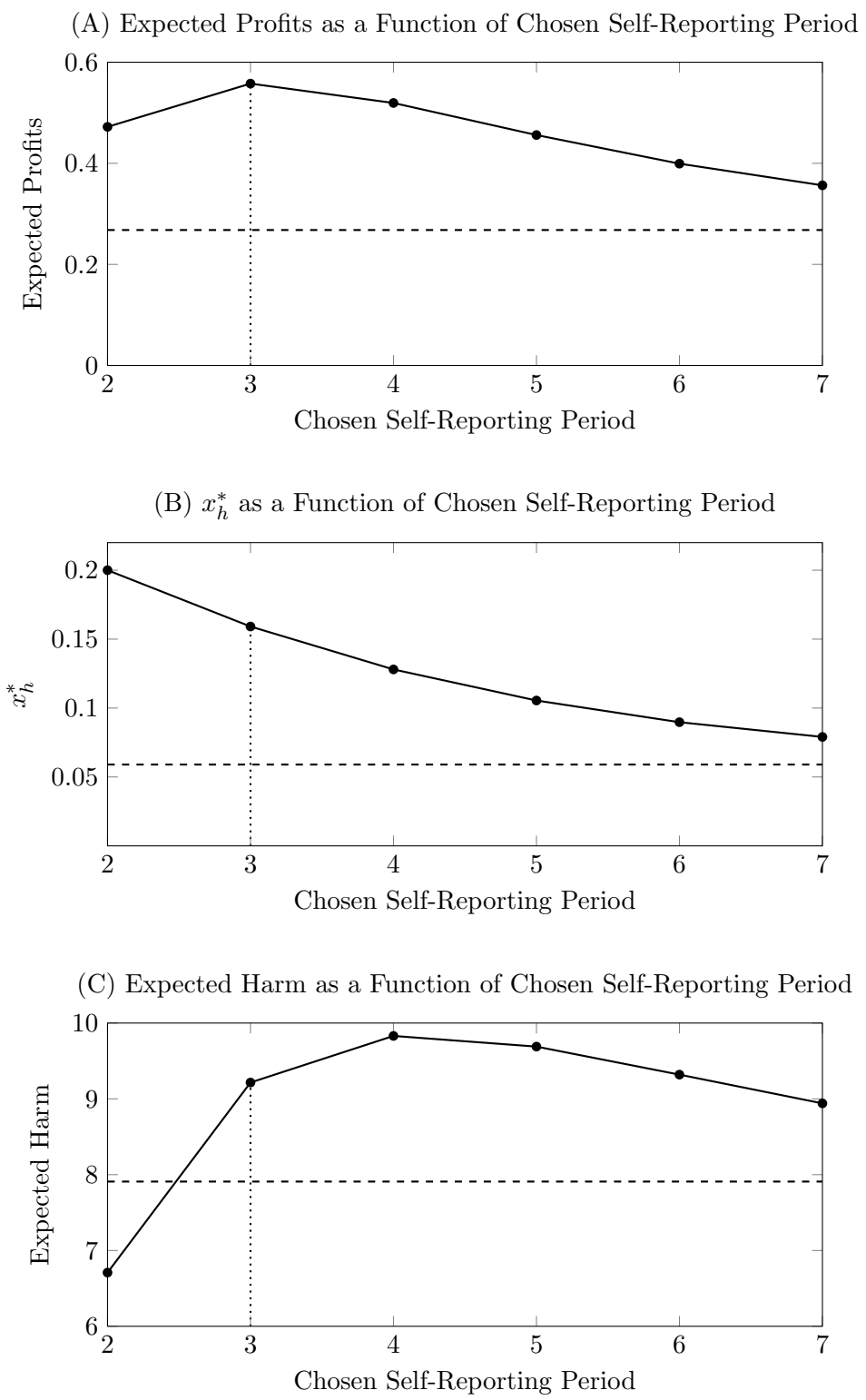


Figure 7: **Self-Reporting.** For the parameters $h = 15$, $\pi_b = 4$, $\pi_h = 3$, $\delta = 0.90$, $\bar{p} = 0.4$, $f_c = 2$, and $D = .5$. Panel (A) shows the corporation's expected profits from choosing a particular self-reporting period, maximized here at 3. Panel (B) shows the corporation's optimal choice of x_h as a function of the choice of self-reporting period. Panel (C) shows the total expected harm as a function of the chosen self-reporting period. In all three graphs, the dashed lines indicate the level when no discounts for self-reporting are offered.

the corporation would choose if no discount were offered for self-reporting so that the corporation will never self-report. The vertical line at 3 indicates the corporation's profit-maximizing choice of self-reporting period and harm.

Finally, Panel (C) of figure 7 shows the expected level of harm associated with each self-reporting strategy. The horizontal dashed line is the expected level of harm if no discounts for self-reporting are offered. Note that the self-reporting program would be beneficial if it meant that the firm only engaged in misconduct in period 1 and self-reported in period 2. The reason is that, although the corporation sets x_h higher, the harm only persists for one period. However, as Panel (A) shows, the optimal strategy is for the corporation to self-report in period 3. Because the harm is expected to persist longer, self-reporting in period 3 increases the expected harm. Harm would be even greater if the firm self-reports in period 4, because of the increased expected duration of harm. However, doing so results in lower profits than self-reporting in period 3 and thus is not the chosen strategy. The expected harm then declines if the corporation were to follow the suboptimal strategy of reporting in periods 5 and onward. These are due to the interplay between the reduction in x_h that comes with self-reporting later that is shown in Panel (B) and the increase in the expected duration of the harm.

Table 5: Self-Reporting

Fine Rate	Self-Reporting Discount	x_h^*	Self-Reporting Period	Profits From Misconduct	Harm
2.0	0%	0.059	NA	0.268	7.909
2.0	50%	0.159	3	0.558	9.216
2.0	90%	0.254	3	0.971	11.636
1.0	0%	0.200	NA	1.026	14.583
1.0	50%	0.230	5	1.111	14.312
1.0	90%	0.278	4	1.293	14.475
0.5	0%	0.281	NA	1.559	17.283
0.5	50%	0.285	8	1.570	17.190
0.5	90%	0.299	6	1.616	17.025

The discussion above shows how a self-reporting regime can increase the harm caused. Table 5 shows the corporation's choice of the level of the harmful activity and self-reporting period as a function of the underlying fine rate (2.0, 1.0, and 0.5) and self-reporting discount (0%, 50%, and 90%). Figure 7 and the discussion above is based on a fine rate of 2.0 and a self-reporting discount of 50%, shown in row two. The third row of Table 5 shows that a self-reporting discount of 90% would cause far more social harm than not offering a discount at all when the $f_c = 2$.

Now consider a fine rate of 0.5. When there is no discount for self-reporting, the first row shows that the corporation engages in the profit-maximizing level of misconduct until misconduct is detected. The second row of the table considers a 50% discount for self-reporting. The lower fine rate (0.5 instead of 2.0) means that a 50% reduction results in lower benefits to the corporation of self-reporting. This means that the corporation waits longer, and does not self-report until period 8. Moreover, because it is unlikely that the corporation will actually make it to period 8 (there is a 97.2% chance that the corporation will be caught before that), the corporation increases x_h, but by much less than when the fine was 2.0. The final column of the table shows that the effect of a 50% self-reporting discount is that the benefits from the decreased expected duration of harm are greater than the costs of increased x_h, so that total expected harm goes down. A 90% discount

would lower it even further.

Table 5 also calls into question the "best" real-world policy choices. Offering discounts when fines are low (0.5 or 1.0 in this case) has a small beneficial effect on reducing social harm. Offering discounts when fines are relatively high (2.0) significantly increases social harm. In the context of this example, self-reporting regimes are only justifiable if fines are set much too low. If that is the case, it would be much better to simply increase fines.

Because discounts for self-reporting work on two margins, predicting their effects in a particular situation seems to require much more knowledge than those designing policies are likely to have. This makes it even more difficult—if not impossible—to say whether a general policy would increase or decrease social harm.

Whereas in the above we considered a multi-period model of misconduct, in some cases a one-period model is sufficient because the misconduct is a "one-off" opportunity available to the corporation for a single period. In this case, we have the following result that, like Proposition 7 is stark.

Proposition 9. *If the misconduct opportunity will only be available in a single period, offering a discount for self-reporting never reduces the harm caused from the misconduct and may in fact increase the harm.*

Proof: See appendix.

The intuition is straightforward. Self-reporting is always voluntary, so corporations will only self-report when the fine following the self-reporting is less than the expected fine from not reporting. If the discount is not large enough to encourage self-reporting it will have no effect. If instead it is large enough to encourage self-reporting, it will lower the expected costs of misconduct, so the corporation will increase the level of the harmful activity.

Kaplow and Shavell (1994) also looks at self-reporting, although in a different context than the one we examine. In the situations they analyze they find that self-reporting can be beneficial because it potentially lowers enforcement costs. We have not explicitly incorporated enforcement costs into our analysis. It is clear that if those costs are significant, they should be accounted for in assessing the *total* welfare implications of various approaches since enforcement costs do reduce welfare. What our analysis shows is that any assumed reductions in enforcement costs must be weighed against the potential for the counterproductive increases in harm that can occur when significant discounts are promised to encourage self-reporting. Since these harms can be significant, it should not be assumed that self-reporting programs can be justified simply on the basis of reducing enforcement costs.[30]

In some cases, the effect of self-reporting by corporations is to move penalties off of the corporation and on to individuals. Indeed, this is the motivation of much of the DOJ's enforcement policies in the US. However, if self-reporting transfers liability onto the manager, the corporation can simply adjust the manager's compensation structure as we showed in Section 3.3. In fact, as we discussed in Section 1.1, it is extremely rare for high-level managers to face liability. Key decision makers in the corporation may not bear any costs when capturing the discount if they are able to "throw under the bus" employees who are purported to be "bad apples," as happened in the Boeing case discussed in section 1.1.

[30]Relatedly, Motta and Polo (2003) show that offering leniency for a cartel member who reports on other cartel members can make enforcement more effective in some cases, but also may foster collusion by lowering the expected cost of misbehavior.

5 Concluding Remarks

This paper starts from the simple observation that corporations and their agents can often benefit financially by doing things that break laws and cause significant harm to others, even after accounting for fines they may be required to pay if their wrongdoing is detected. Fines and other penalties may be considered just a "cost of doing business."

We have shown that when law enforcement is weak, standard corporate governance practices may make deterrence even less effective, since corporations can proactively use equity-based managerial contracts, insurance, and indemnification arrangements to offset the deterrent impact of fines. We also showed that providing incentives for corporations to cooperate with enforcement authorities by offering significant reductions in penalties for establishing voluntary compliance programs or self reporting misconduct may increase the amount of harm corporations cause. Whereas cooperation can lead to earlier detection, the reduced fines make misconduct prior to detection more profitable and generally lead to more harm before detection. We have the stark result that if the choice of misconduct is available only once, these programs will never decrease harm and may actually increase harm. One implication of our analysis is that when enforcement is ineffective, making changes that might be viewed as improvements, such as establishing voluntary compliance programs, may in fact make matters worse.

Whereas our paper has shown that standard corporate governance practices can undermine law enforcement, our results do not imply that all less-than-perfect law enforcement policies will be ineffective or counterproductive. We showed that even inadequate fines on corporations can reduce the level of the harmful activity, but we also showed there are cases where fines have no effect and cases where well-intentioned policy changes such as basing fines on the harm done rather than profits gained can be counterproductive.

Our analysis suggests many directions for future research. First, we have been primarily focused on how fines affect the financial incentives of corporations and their managers to engage in wrongdoing. Although corporations cannot be jailed, managers can. Jed Rakoff, a prominent Senior Judge in the Southern District of New York, argues that the most effective way to prevent serious misconduct by corporations is to have leaders responsible for misconduct face a genuine possibility of prison time. More research needs to be done on how increasing the chance that culpable managers will go to jail can be used in addition to other sanctions to help prevent social harm. A possible approach would be to reduce the extremely high bar for criminal conviction for individuals classified as "responsible officers" in a corporation based on the principle that power within a corporation must come with accountability.[31]

Another set of issues that should be studied concerns how law enforcement directed at corporations interacts with employment issues. Low-level employees may be given incentives to break laws on behalf of the corporation. If they engage in misconduct, they may suffer consequences when the misconduct is revealed (potentially being labeled "bad apples"), but if they do not engage, they may be punished by their superiors for not performing well. The case of the fraudulent account-opening scandal at Wells Fargo illustrates this problem. Since our model has only one manager, it cannot be used to explore how hierarchies within large corporations affect the severity of corporate misconduct and the additional challenges they might pose in deterrence. It often seems that organizations are designed to protect those higher in the hierarchy, which no doubt undermines deterrence since those at higher levels are the major decision makers.

[31] For example, a responsible officer might be convicted if it is proved beyond a reasonable doubt that the individual effectively knew but was "willfully blind" by refusing to acknowledge, persistent misconduct that, was able to take action to prevent it, and instead failed to take such action and thus enable the misconduct.

Other important topics worthy of further research relate to information flows and disclosures, which are important for both preventing and detecting corporate misconduct. Examples include the design of monitoring and reporting structures within corporations, including internal regimes that govern disclosures to other stakeholders, whistleblower programs, and the use (and potential abuse) of non-disclosure agreements that can interfere with the detection of misconduct. Whistleblower policies that protect and encourage those who can provide critical information about corporate misconduct to outsiders are extremely important and raise many issues worthy of future analysis.[32] Many practices such delegating investigation to the offending corporation and deferred prosecution agreements raise issues around what information is produced and shared and often hide and obscure information that is highly valuable for society, such as information about dangerous products.[33]

As mentioned in the introduction, corporate fines may be insufficient to deter corporate misconduct because corporations can borrow excessively and use the bankruptcy process strategically to limit their legal liability, including from misconduct. Tax subsidies that encourage excessive borrowing by corporations, and the treatment of prior misconduct in bankruptcy processes should be re-examined so as to address the potential distortions and social harm they may create in the context of misconduct.

As mentioned in the introduction, corporate fines may be insufficient to deter corporate misconduct because corporations can borrow excessively and use the bankruptcy process strategically to limit their legal liability, including from misconduct. Tax subsidies that encourage excessive borrowing by corporations, and the treatment of prior misconduct in bankruptcy processes should be re-examined so as to address the potential distortions and social harm they may create in the context of misconduct.

Another reason we noted for why fines may be too low to be effective deterrent of corporate misconduct involves concerns that imposing large fines may lead to collateral harm for innocent stakeholders such as employees. This issue can be addressed if enforcers require that the corporation issue new equity to cover fines. Doing so ensures that the fine will affect shareholders and managers who benefit from profitable misconduct without depleting corporate resources in ways that harm innocent stakeholders. (Atkinson, 2023*b*). More generally, it is important to study ways that sanctions can be designed so they can best deter corporate misconduct but avoid collateral harm and other adverse consequences. The specific approaches may depend on the type of misconduct.

Requiring changes to internal governance mechanisms is another possible approach to deterrence of corporate misconduct. For example, the board of directors plays a central role in internal governance, and it has a duty to ensure the corporation obeys the law. Finding ways to create incentives for the board to fulfill its duties in this regard is worth further study.[34] More broadly, it would be useful to develop and evaluate mechanisms and policies that change internal corporate governance and its interaction with external governance mechanisms in various contexts.

[32]On some issues related to whistleblower policies and their potential, see Baer (2016) and Jennings (2025). Dyck, Morse and Zingales (2010) show that whistleblowers play an important role in uncovering corporate fraud, and Buell (2024) recommends more focus on whistleblowers for earlier detection of corporate wrongdoing. The role of media in exposing misconduct is also an area for future research.

[33]On the lack of potentially important information in legal settlements see, for example, Burch and Lahav (2020) and Baer (2020).

[34]For example, Pfleiderer (2020) shows how misconduct could be better deterred by requiring, especially in cases of recidivist firms, that the board of directors include director who substantially benefits when the firm does not incur fines for misconduct but receives no benefit from increases in corporate profits. This would result in a director with high-powered incentives to expose and halt corporate misconduct, which would enhance external enforcement.

Requiring changes to internal governance mechanisms is another possible approach to deterrence of corporate misconduct. For example, the board of directors plays a central role in internal governance, and it has a duty to ensure the corporation obeys the law. Finding ways to create incentives for the board to fulfill its duties in this regard is worth further study.[35] More broadly, it would be useful to develop and evaluate mechanisms and policies that change internal corporate governance and its interaction with external governance mechanisms in various contexts.

In our analysis we implicitly assumed that laws are appropriately written to protect the public from harm and we focused on problems in their enforcement. It is well understood that the realities of political economy often lead to laws that are not necessarily in the public interest. Corporations can shape the ways laws are written through lobbying and use their influence to weaken enforcement mechanisms so that the process tolerates more harmful activities. A better understanding of how corporate governance affects social welfare clearly requires that social scientists and economists study the political economy aspects of these issues.

Our focus has been on law enforcement in situations where the interests of the corporation, its shareholders and its managers are aligned with each other but in conflict with society's interests. In other words, we are assuming that shareholder governance works well, but external governance fails. There may well be situations in which managers take actions for their own benefit that harm both the corporation's shareholders and society due to a combined failure of internal and external governance. One issue is that the horizon of managers is typically shorter than that of the corporation and its shareholders. Managers can capture near-term benefits from misconduct (e.g., higher bonuses) and avoid accountability for the harms they caused to both the shareholders and society that become evident later after they leave the corporation. In these cases improvements in internal governance that better align manager's incentives with those of the shareholders can reduce social harm. Of course, as we have shown, whenever misconduct benefits shareholders, improvements in internal governance can increase social harm.

Finally, it is important to note that issues around laws and their enforcement are complicated by the fact that laws interact with one another within and across jurisdictions, and corporations and economic activities often span multiple jurisdictions within and across countries. These considerations lead to additional important issues and research topics. For example, corruption and poor rule of law in one nation can be exploited by corporations domiciled in other jurisdictions. Effective law enforcement may require cooperation across authorities and jurisdictions, but governments that can create and enforce laws to prevent harms away from occurring outside their jurisdictions might lack the incentives to do so. Money laundering, which enables kleptocracy and many other serious crimes, is a clear example of this problem. Multinational corporations may also benefit from the support of their "home" governments in resisting law enforcement elsewhere.

Powerful financial incentives can motivate corporations to create huge benefits for society. Realizing these benefits requires that corporations and managers conform to the basic rules of society, which Milton Friedman argues is their responsibility. When law enforcement is weak and managers have strong incentives to "make as much money as possible," this condition is often not met. Our results emphasize the need for law enforcement policies that are better designed to account for the powerful incentives driving corporate decision-making.

[35] For example, Pfleiderer (2020) shows how misconduct could be better deterred by requiring, especially in cases of recidivist firms, that the board of directors include director who substantially benefits when the firm does not incur fines for misconduct but receives no benefit from increases in corporate profits. This would result in a director with high-powered incentives to expose and halt corporate misconduct, which would enhance external enforcement.

References

Admati, Anat, and Martin Hellwig. 2024. *The Bankers' New Clothes: What's Wrong with Banking and What to Do About it.* Princeton University Press.

Admati, Anat R. 2017. "a Skeptical View of Financialized Corporate Governance." *Journal of Economic Perspectives*, 31(3): 131–150.

Admati, Anat R. 2021. "Capitalism, Laws, and the Need for Trustworthy Institutions." *Oxford Review of Economic Policy*, 37(4): 678–689.

Admati, Anat R. 2025. "Unjust Debts: How Our Bankruptcy System Makes America more Unequal by Melissa B. Jacoby." *Journal of Economic Literature*.

Admati, Anat R, Peter M DeMarzo, Martin F Hellwig, and Paul Pfleiderer. 2018. "The Leverage Ratchet Effect." *The Journal of Finance*, 73(1): 145–198.

Alexander, Cindy R, and Mark A Cohen. 2015. "The Evolution of Corporate Criminal Settlements: An Empirical Perspective on Non-Prosection, Deferred Prosecution, and Plea Agreements." *Am. Crim. L. Rev.*, 52: 537.

Argenton, Cédric, and Eric van Damme. 2014. "Optimal Deterrence of Illegal Behavior under Imperfect Corporate Governance."

Arlen, Jennifer. 1994. "The Potentially Perverse Effects of Corporate Criminal Liability." *The Journal of Legal Studies*, 23(2): 833–867.

Arlen, Jennifer, and Reinier Kraakman. 1997. "Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes." *NYUL Rev.*, 72: 687.

Atkinson, Nathan. 2020. "If Not the Index Funds, Then Who?" *Berkeley Bus. LJ*, 17: 44.

Atkinson, Nathan. 2023*a*. "Corporate Liability, Collateral Consequences, and Capital Structure." *Colum. Bus. L. Rev.*, 1.

Atkinson, Nathan. 2023*b*. "Mandatory Equity Issuances as a First-Best Solution to Punishing Corporate Misconduct." *Univ. of Wisconsin Legal Studies Research Paper*, , (1791).

Atkinson, Nathan. 2023*c*. "Profiting from Pollution." *Yale Journal on Regulation Bulletin*, 41: 1.

Badawi, Adam B, and Frank Partnoy. 2022. "Social Good and Litigation Risk." *Harv. Bus. L. Rev.*, 12: 315.

Baer, Miriam H. 2016. "Reconceptualizing the Whistleblower's Dilemma." *UCDL Rev.*, 50: 2215.

Baer, Miriam H. 2020. "The Information Shortfalls of Prosecuting Irresponsible Executives." *DePaul L. Rev.*, 70: 191.

Baker, Andrew C, David F Larcker, Charles G McClure, Durgesh Saraph, and Edward M Watts. 2024. "Diversity Washing." *Journal of Accounting Research*, 62(5): 1661–1709.

Baker, Tom, and Sean J Griffith. 2019. *Ensuring corporate misconduct: How liability insurance undermines shareholder litigation.* University of Chicago Press.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169–217.

Black, Bernard, Brian Cheffins, and Michael Klausner. 2005. "Outside Director Liability." *Stan. L. Rev.*, 58: 1055.

Blunt, Katherine. 2022. *California Burning: The Fall of Pacific Gas and Electric–and What it Means for America's Power Grid.* Penguin.

Bratton, William W, and Michael L Wachter. 2011. "The political economy of fraud on the market." *U. Pa. L. Rev.*, 160: 69.

Buell, Samuel W. 2024. "The Limits of Individual Prosecutions in Deterring Corporate Fraud." *Wake Forest Law Review*, 59(3).

Burch, Elizabeth Chamblee, and Alexandra D Lahav. 2020. "Information for the Common Good in Mass Torts." *DePaul L. Rev.*, 70: 345.

Byars, Kaleb. 2025. "Recidivist Organizational Offenders and the Organizational Sentencing Guidelines." *Boston College Law Review*, 66(3).

Chen, Kong-Pin, and CY Cyrus Chu. 2005. "Internal Control Versus External Manipulation: A Model of Corporate Income Tax Evasion." *rand Journal of Economics*, 151–164.

Coffee, John C. 2022. "Nosedive: Boeing and the Corruption of the Deferred Prosecution Agreement." *Available at SSRN 4105514*.

Coffee Jr, John C. 2020. *Corporate Crime and Punishment: The Crisis of Underenforcement.* Berrett-Koehler Publishers.

Copeland, Katrice Bridges. 2014. "The Crime of Being in Charge: Executive Culpability and Collateral Consequences." *Am. Crim. L. Rev.*, 51: 799.

Copeland, Katrice Bridges. 2016. "The Yates Memo: Looking for Individual Accountability in all the Wrong Places." *Iowa L. Rev.*, 102: 1897.

Crocker, Keith J, and Joel Slemrod. 2005. "Corporate tax evasion with agency costs." *Journal of public Economics*, 89(9-10): 1593–1610.

DeMarzo, Peter M., Michael J. Fishman, and Kathleen M. Hagerty. 2005. "Self-Regulation and Government Oversight." *The Review of Economic Studies*, 72(3): 687–706.

Deng, Sheran, Eitan Goldman, and Matthias Kahl. 2021. "Regulating a Firm with Agency Problems." *Kelley School of Business Research Paper*, , (2020-64).

Desai, Mihir A, Alexander Dyck, and Luigi Zingales. 2004. "Corporate Governance and Taxation." 75, bepress.

Desai, Mihir A, and Dhammika Dharmapala. 2006. "Corporate Tax Avoidance and High-Powered Incentives." *Journal of financial Economics*, 79(1): 145–179.

Döttling, Robin J, Doron Y Levit, Nadya Malenko, and Magdalena A Rola-Janicka. 2024. "Voting on Public Goods: Citizens vs. Shareholders." National Bureau of Economic Research.

Dyck, Alexander, Adair Morse, and Luigi Zingales. 2010. "Who Blows the Whistle on Corporate Fraud?" *The journal of finance*, 65(6): 2213–2253.

Dyck, Alexander, Adair Morse, and Luigi Zingales. 2024. "How Pervasive is Corporate Fraud?" *Review of Accounting Studies*, 29(1): 736–769.

Efendi, Jap, Anup Srivastava, and Edward P Swanson. 2007. "Why Do Corporate Managers Misstate Financial Statements? the Role of Option Compensation and Other Factors." *Journal of financial economics*, 85(3): 667–708.

Egan, Mark, Gregor Matvos, and Amit Seru. 2019. "The Market for Financial Adviser Misconduct." *Journal of Political Economy*, 127(1): 233–295.

Eisinger, Jesse. 2017. *The Chickenshit Club: Why the Justice Department Fails to Prosecute Executives.* Simon and Schuster.

Estlund, Cynthia. 2017. "The Black Hole of Mandatory Arbitration." *NCL Rev.*, 96: 679.

Ewing, Jack. 2017. *Faster, Higher, Farther: How One of the World's Largest Automakers Committed a Massive and Stunning Fraud.* WW Norton & Company.

Findley, Michael G, Daniel L Nielson, and JC Sharman. 2024. "Banking Bad? a Global Field Experiment on Risk, Reward, and Regulation." *American Journal of Political Science.*

Garrett, Brandon L. 2014. *Too Big to Jail: How Prosecutors Compromise with Corporations.* Harvard University Press.

Garrett, Brandon L. 2020. "Declining Corporate Prosecutions." *Am. Crim. L. Rev.*, 57: 109.

Hart, Oliver, and Luigi Zingales. 2017. "Companies Should Maximize Shareholder Welfare not Market Value." *ECGI-Finance Working Paper*, , (521).

Jacoby, Melissa B. 2024. *Unjust Debts: How Our Bankruptcy System Makes America more Unequal.* Princeton University Press.

Jennings, Andrew. 2025. "Early Whistleblowing." *63 American Criminal Law Review (forthcoming 2026), Available at SSRN:5262641.*

Kaplow, Louis, and Steven Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior." *Journal of Political Economy*, 102(3): 583–606.

Karpoff, Jonathan M, D Scott Lee, and Valaria P Vendrzyk. 1999. "Defense Procurement Fraud, Penalties, and Contractor Influence." *Journal of Political Economy*, 107(4): 809–842.

Khan, Mozaffar, Suraj Srinivasan, and Liang Tan. 2017. "Institutional Ownership and Corporate Tax Avoidance: New Evidence." *The Accounting Review*, 92(2): pp. 101–122.

Kornhauser, Lewis A. 1982. "An Economic Analysis of the Choice Between Enterprise and Personal Liability for Accidents." *Calif. L. Rev.*, 70: 1345.

Levi, Michael, and Peter Reuter. 2006. "Money Laundering." *Crime and Justice*, 34(1): 289–375.

Lund, Dorothy S, and Elizabeth Pollman. 2021. "The Corporate Governance Machine." *Colum. L. Rev.*, 121: 2563.

McAfee, R Preston, Hugo M Mialon, and Sue H Mialon. 2008. "Private v. Public Antitrust Enforcement: a Strategic Analysis." *Journal of Public Economics*, 92(10-11): 1863–1875.

Motta, Massimo, and Michele Polo. 2003. "Leniency programs and cartel prosecution." *International journal of industrial organization*, 21(3): 347–379.

Mullin, Wallace P, and Christopher M Snyder. 2010. "Should firms be allowed to indemnify their employees for sanctions?" *The Journal of Law, Economics, & Organization*, 26(1): 30–53.

Newman, Harry A, and David W Wright. 1990. "Strict Liability in a Principal-Agent Model." *International Review of Law and Economics*, 10(3): 219–231.

Oded, Sharon. 2016. "Coughing Up Executives or Rolling the Dice: Individual Accountability for Corporate Corruption." *Yale L. & Pol'y Rev.*, 35: 49.

Pfleiderer, Paul. 2020. "The Milton Friedman Constraint: a Proposal for Improving Corporate Governance." *Available at SSRN 3757008*.

Podgor, Ellen S. 2012. "Overcriminalization: New approaches to a growing problem." *J. Crim. L. & Criminology*, 102: 529.

Podgor, Ellen S. 2018. "White Collar Shortcuts." *U. Ill. L. Rev.*, 925.

Polinsky, A Mitchell. 1980. "Private Versus Public Enforcement of Fines." *The Journal of Legal Studies*, 9(1): 105–127.

Polinsky, A Mitchell, and Steven Shavell. 1993. "Should Employees be Subject to Fines and Imprisonment Given the Existence of Corporate Liability?" *International Review of Law and Economics*, 13(3): 239–257.

Polinsky, A Mitchell, and Steven Shavell. 2007. "The Theory of Public Enforcement of Law." *Handbook of law and economics*, 1: 403–454.

Pollman, Elizabeth. 2021. "The Supreme Court and the Pro-Business Paradox, 135 Harv." *L. Rev*, 220: 247–61.

Raghunandan, Aneesh. 2021. "Financial Misconduct and Employee Mistreatment: Evidence from Wage Theft." *Review of Accounting Studies*, 26(3): 867–905.

Rakoff, Judge Jed S. 2021. *Why the Innocent Plead Guilty and the Guilty Go Free: And Other Paradoxes of Our Broken Legal System.* Farrar, Straus and Giroux.

Robison, Peter. 2022. *Flying Blind: The 737 Max Tragedy and the Fall of Boeing.* Anchor.

Rudman, William J, John S Eberhardt, William Pierce, and Susan Hart-Hester. 2009. "Healthcare Fraud and Abuse." *Perspectives in Health Information Management/AHIMA, American Health Information Management Association*, 6(Fall).

Schwab, Stewart J, and Randall S Thomas. 2006. "An Empirical Analysis of Ceo Employment Contracts: What Do Top Executives Bargain for." *Wash. & Lee L. Rev.*, 63: 231.

Shapira, Roy. 2024. "Conceptualizing Caremark."

Shapira, Roy, and Luigi Zingales. 2023. "Is Pollution Value-Maximizing? The Dupont Case." *European Corporate Governance Institute-Law Working Paper*.

Taub, Jennifer. 2020. *Big Dirty Money: The Shocking Injustice and Unseen Cost of White Collar Crime*. Penguin.

Thomas, W Robert. 2021. "Corporate Criminal Law is Too Broad-Worse, It's Too Narrow." *Ariz. St. LJ*, 53: 199.

Welch Jr, James. 2019. "The Volkswagen Recovery: Leaving Scandal in the Dust." *Journal of Business Strategy*, 40.

Zingales, Luigi. 2015. "Presidential Address: Does Finance Benefit Society?" *The Journal of Finance*, 70(4): 1327–1363.

Appendix: Proofs of the Propositions

Proof of Proposition 1: The total profit maximization problem is

$$\max_{x_b,x_h} (\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}) \quad \text{s.t.} \quad x_b + x_h \leq 1, \tag{A1}$$

and the profit maximizing levels are:

$$x_b^* = \frac{\pi_b^2}{\pi_b^2 + \pi_h^2}; \quad x_h^* = \frac{\pi_h^2}{\pi_b^2 + \pi_h^2}. \tag{A2}$$

Given compensation $Comp = S + \alpha \times \left(\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}\right)$, the manager's utility maximization problem is:

$$\max_{x_b,x_h} (S + \alpha \times (\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}) - \beta \times (x_b + x_h)) \quad \text{s.t.} \quad x_b + x_h \leq 1. \tag{A3}$$

If α, the manager's share of total profits, is set equal to $2\beta/\sqrt{\pi_b^2 + \pi_h^2}$, the manager's optimization problem becomes:

$$\max_{x_b,x_h} \left((1-\alpha)S + \frac{2\beta}{\sqrt{\pi_b^2 + \pi_h^2}} \times (\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}) - \beta \times (x_b + x_h) \right) \quad \text{s.t. } x_b + x_h \leq 1. \tag{A4}$$

The first order conditions for an optimum are:

$$\frac{\beta\pi_b}{\sqrt{\pi_b^2 + \pi_h^2}} \times \frac{1}{\sqrt{x_b}} - \beta = 0 \implies x_b = \frac{\pi_b^2}{\pi_b^2 + \pi_h^2} = x_b^* \tag{A5}$$

$$\frac{\beta\pi_h}{\sqrt{\pi_b^2 + \pi_h^2}} \times \frac{1}{\sqrt{x_h}} - \beta = 0 \implies x_h = \frac{\pi_h^2}{\pi_b^2 + \pi_h^2} = x_h^* \tag{A6}$$

and so the manager's choices for x_b and x_h maximize the total corporate profits from the activities.

The proposition also holds if the maximization problem is to maximize profits *net* of what the manager receives through the profit sharing defined in his compensation contact as long as $\alpha < 1/2$. This is because shareholders' share of profits will be equal to:

$$(1-\alpha)\left(\frac{\alpha\pi_b^2}{2\beta} + \frac{\alpha\pi_h^2}{2\beta}\right), \tag{A7}$$

which is increasing in α if $\alpha < 1/2$. This will hold given our assumption that $\beta < \sqrt{\pi_b^2 + \pi_h^2}/4$. ■

Proof of Proposition 2: Let $\Pi^* = \pi_b\sqrt{x_b^*} + \pi_h\sqrt{x_h^*}$ where x_b^* and x_h^* are given in (A2) and are the values that maximize total profit when there are no fines. Let $\Pi^\psi = \pi_b\sqrt{1-\psi} + \pi_h\sqrt{\psi}$, which is equal to the profits realized when the harmful activity is set at the threshold that does not lead to a fine and the level of the beneficial activity is set equal to $1-\Upsilon$. There are three cases to consider:

Case 1: $x_h^* \leq \psi$

Case 2: $x_h^* > \psi$ and $\Pi^* - \Pi^\psi > F_c$

Case 3: $x_h^* > \psi$ and $\Pi^* - \Pi^\psi \leq F_c$

In case 1 the optimal level of the harmful activity falls below the threshold for a fine so the corporation will obviously choose that level. In case 2, the optimal level is above the threshold for a fine, but the corporation makes higher profits even after paying the fine than what it would have made if it set $x_h = \psi$ and avoided the fine. So again the corporation will choose the same level of the harmful activity as it would have chosen without a fine. In case 3 the fine is large enough to make $\Pi^* - F_c$, the maximum profits that can be attained by setting $x_h = x_h^*$ and paying the fine, smaller than the profit that can be attained when $x_h = \psi$ and no fine is imposed.

To give the manager incentives to do what maximizes profits, it is clear that in cases 1 and 2 the same compensation contract can be used as when there are no fines levied, i.e,. *Comp* is equal to S plus α times corporate profits (offset by any payment of a fine) where $\alpha = 2\beta/\sqrt{\pi_b^2 + \pi_h^2}$. (It may be necessary to adjust the value of S in case 2 to ensure that the manager's compensation is worth at least his best outside opportunity since profits are diminished by the fine.) In case 3 the compensation contract must be adjusted so the manager has the incentive to set $x_b = 1 - \psi$ since the harmful activity will be capped at $x_h = \psi$.

This observation implies that α must be set so that the following first order condition for a maximum holds:

$$\frac{\alpha\pi_b}{2\sqrt{1-\psi}} - \beta = 0 \quad \Rightarrow \quad \alpha = \frac{2\beta\sqrt{1-\psi}}{\pi_b}. \tag{A8}$$

Thus in every case the manager can be motivated to maximize profits. ■

Proof of Proposition 3: The fine transforms the corporation's profit maximization problem into:

$$\max_{x_b, x_h} \left(\pi_b\sqrt{x_b} + (\pi_h - f_c)\sqrt{x_h}\right) \qquad \text{s.t.} \qquad x_b + x_h \leq 1. \tag{A9}$$

For any level of fine the solution is

$$x_b^* = \frac{\pi_b^2}{\pi_b^2 + \left(\max\left(\pi_h - f_c, 0\right)\right)^2}; \quad x_h^* = \frac{\left(\max\left(\pi_h - f_c, 0\right)\right)^2}{\pi_b^2 + \left(\max\left(\pi_h - f_c, 0\right)\right)^2} \tag{A10}$$

For $f_c \geq \pi_h$, we have $x_b^* = 1$ and $x_h^* = 0$. Otherwise, we have $x_h^* > 0$ and the harmful activity is not fully deterred. If the fraction of the profits the manager receives in his compensation is adjusted to

$$\alpha = \frac{2\beta}{\sqrt{\pi_b^2 + \left(\max\left(\pi_h - f_c, 0\right)\right)^2}}. \tag{A11}$$

the manager will have the incentive to choose the optimal levels in (A10). ■

Proof of Proposition 4: The manager's maximization problem when facing the fixed fine is:

$$\max_{x_b, x_h} \left(\alpha \times \left(\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}\right) - F_m \times I_{x_h > \psi} - \beta \times (x_b + x_h)\right)$$
$$\text{s.t. } x_b + x_h \leq 1 \tag{A12}$$

where $I_{x_h > \psi}$ is the indicator function, which is equal to 1 when $x_h > \psi$ and is 0 otherwise. First, assume that $\alpha = \bar{\alpha}$, and let $x_b^* = \pi_b^2/\left(\pi_b^2 + \pi_h^2\right)$ and $x_h^* = \pi_h^2/\left(\pi_b^2 + \pi_h^2\right)$, which are the optimal choices when there is no fine. Finally, let

$$U^* = \bar{\alpha} \times \left(\pi_b\sqrt{x_b^*} + \pi_h\sqrt{x_h^*}\right) - F_m \times I_{x_h^* > \psi} - \beta \times (x_b^* + x_h^*) \tag{A13}$$
$$U^\psi = \bar{\alpha} \times \left(\pi_b\sqrt{x_b^*} + \pi_h\sqrt{\psi}\right) - \beta \times (x_b^* + \psi) \tag{A14}$$

where $\bar{\alpha} = 2\beta/\sqrt{\pi_b^2 + \pi_h^2}$. U^* is the manager's utility when the activity levels are set equal to the optimal levels without a fine and a fine is paid if $x_h^* > \psi$. U^ψ is the manager's utility when $x_b = x_b^*$ and x_h is set equal to ψ, so that no fine is incurred. (Note that when $x_h = \psi$, it is still optimal for the manager to set x_b equal to x_b^*.) As with the fixed fine on the corporation, we have three cases to consider.

Case 1: $x_h^* \leq \psi$

Case 2: $x_h^* > \psi$ but $U^* \geq U^\psi$

Case 3: $x_h^* > \psi$ and $U^* < U^\psi$

In both case 1 and case 2, the fixed fine has no effect on the manager's choice of the harmful activity level. Only in case 3 is there potentially an effect. When case 3 occurs, it is possible for the corporation to set $\alpha > \bar{\alpha}$ so that the manager will chose $x_h = x_h^*$ instead of $x_h^* = \psi$. The minimum level of α that will change the manager's choice of x_h from x_h^* to ψ, call it α^{adj}, solves

the following equation:

$$\alpha^{adj} \times \left(\pi_b\sqrt{x_b^*} + \pi_h\sqrt{x_h^*}\right) - F_m - \beta \times (x_b^* + x_h^*) = \alpha^{adj} \times \left(\pi_b\sqrt{x_b^*} + \pi_h\sqrt{\psi}\right) - \beta \times (x_b^* + \psi)$$

$$\alpha^{adj} = \frac{F_m + \beta \times (x_h^* - \psi)}{\pi_h\left(\sqrt{x_h^*} - \sqrt{\psi}\right)}. \tag{A15}$$

Note that whenever Case 3 occurs, $\alpha^{adj} > 2\beta/\sqrt{\pi_b^2 + \pi_h^2} = \bar{\alpha}$. Since α^{adj} must be less than 1, this requires that $F_m \leq \pi_h\left(\sqrt{x_h^*} - \sqrt{\psi}\right) - \beta \times (x_h^* - \psi)$. If the level of the fine exceeds that, it cannot be nullified by a change in the level of α. In that case it is easy to see that x_h will be set equal to ψ. . ■

Proof of Proposition 5: With the existence of the fine, the manager's profit maximization problem becomes:

$$\max_{x_b, x_h} \left(\alpha \times (\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}) - f_m\sqrt{x_h} - \beta \times (x_b + x_h)\right) \quad \text{s.t.} \quad x_b + x_h \leq 1 \tag{A16}$$

The first order conditions for the maximization problem and the maximizing values of x_b and x_h are:

$$\begin{aligned} \frac{\alpha\pi_b}{2\sqrt{x_b}} - \beta = 0 \quad &\Rightarrow \quad x_b^* = \left(\frac{\alpha\pi_b}{2\beta}\right)^2 \\ \frac{\max(\alpha\pi_h - f_m, 0)}{2\sqrt{x_h}} - \beta = 0 \quad &\Rightarrow \quad x_h^* = \left(\frac{\max(\alpha\pi_h - f_m, 0)}{2\beta}\right)^2 \end{aligned} \tag{A17}$$

The corporation maximizes profits when α is set so that $x_b + x_h = 1$. This requires that

$$\alpha^{\text{adjusted}} = \frac{\pi_h f_m + \left(4\beta^2\left(\pi_b^2 + \pi_h^2\right) - \pi_b^2 f_m^2\right)^{1/2}}{\pi_b^2 + \pi_h^2} \tag{A18}$$

This solution for α^{adjusted} only applies if it satisfies $\alpha^{\text{adjusted}} \times \pi_h > f_m$. When this is violated, the optimal solution is for the corporation is to set $\alpha = 2\beta/\pi_b$, in which case the manager will choose $x_b = 1$ and $x_h = 0$.

Over the range where $\alpha^{\text{adjusted}} \times \pi_h > f_m$, α^{adjusted} will be higher than the α that the corporation sets when there is no fine, which means that the adjustment leads to more harmful

activity than would be the case without the adjustment, since as can be seen in (A17), x_h^* is increasing in α. This means that the fine causes an adjustment in the manager's compensation that increases the level of the harmful activity beyond what it would be if no adjustment were made. ■

Proof of Proposition 7: Because the corporation voluntarily decides whether and how much to invest in a compliance program, the corporation in effect solves:

$$\max_{x_b,\, x_h,\, p} \; (\pi_b\sqrt{x_b} + \pi_h\sqrt{x_h}) \; - \; p \times (1 - D(p)) \times f_c\sqrt{x_h} \qquad \text{s.t.} \quad x_b + x_h \leq 1, \; p \geq \bar{p}. \tag{A19}$$

Just as before, the manager's compensation contract *Comp* is set to induce the activity levels (x_b^*, x_h^*) the corporation finds optimal.

Of course, a higher probability of detection p coupled with the original fine f_c would more effectively deter misconduct. However, the discount $D(p)$ lowers the fine imposed on the corporation, and thus it may the corporation's *expected* cost of misconduct. Define:

$$\Phi(p) \equiv p\big[1 - D(p)\big] \tag{A20}$$

to be the *expected fine factor*, since taking into account of the discount the expected fine will be $\Phi(p) \times f_c$. For any given $x_h > 0$, the corporation maximizes (A19) by minimizing the expected fine factor, $\Phi(p)$. This leads to two possibilities.

Case 1. Suppose that $\min_p \Phi(p) = \bar{p}$. That is, the expected fine factor for any possible level of p is never lower than $\bar{p}$, the exogenous probability of detection. This means that the discount received from choosing p is not large enough to offset the increased probability of detection, which in turn implies that

$$\pi_h - f_c\Phi(p) \leq \pi_h - \bar{p}f_c, \text{ for all } p \in [\bar{p}, 1]. \tag{A21}$$

In this situation, the adoption of a compliance program weakly decreases the net return from the harmful activity. The corporation will therefore not adopt a compliance program in this situation, and the probability of detection will remain at $\bar{p}$.

Case 2. Suppose $\Phi(p)$ is minimized at $p^\star > \bar{p}$. In this case the compliance program is effective in providing an incentive for the corporation to take actions that increase the probability of detection. The expected fine factor is less than $\bar{p}$ because the discount in the fine more than outweighs the increased probability of detection. In the end this *increases* the profitability of the

harmful activity since

$$\pi_h - \Phi(p^\star) \times f_c > \pi_h - \bar{p} \times f_c. \tag{A22}$$

Thus in case 2 the voluntary compliance program *decreases* social welfare. ■

Proof of Proposition 9: If the government does not use self-reporting and $D = 0$, then the corporation chooses:

$$x_h^* = \frac{(\pi_h - f_c)^2}{\pi_b^2 + (\pi_h - f_c)^2} \tag{A23}$$

Now consider if the government attempts to reduce harm by choosing some $D > 0$. Consider three cases:

1. $D = 1 - p$. In this case, there will be no effect because the expected fine from self-reporting is the same as the expected fine from not self-reporting. Harm remains unchanged.

2. $D < 1 - p$. In this case the corporation will not self-report because it increases the expected fine.

3. $D > 1 - p$. In this case the corporation will self-report. Moreover, because the effective fine is reduced by $D \times f_c\sqrt{x_h}$, the corporation will increase x_h. ■