CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 477

Whistleblower Rewards, False Reports, and Corporate Fraud

Paolo Buccirossi, Giovanni Immordino, and Giancarlo Spagnolo

June 2017
This version September 2017

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unisa.it
ISSN: 2240-9696

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 477

Whistleblower Rewards, False Reports, and Corporate Fraud

Paolo Buccirossi[♥], Giovanni Immordino[♦], and Giancarlo Spagnolo[♣]

Abstract
It is often claimed that rewards for whistleblowers lead to fraudulent reports, but for several US programs this has not been a major problem. We model the interaction be- tween rewards for whistleblowers, sanctions against fraudulent reporting, judicial errors and standards of proof in the court case on a whistleblower.s allegations and the pos- sible follow-up for fraudulent allegations. Balancing whistleblower rewards, sanctions against fraudulent reports, and courts.standards of proof is essential for these policies to succeed. When the risk of retaliation is severe, larger rewards are needed and so are tougher sanctions against fraudulent reports. The precision of the legal system must be su¢ ciently high, hence these programs are not viable in weak institution environ- ments, where protection is imperfect and court precision low, or where sanctions against false reporting are mild. Internal reporting channels may interfere with external ones in unexpected ways.

Keywords: Whistleblowers rewards, False allegations, Judicial errors, Standard of proof, Corporate fraud.

Acknowledgement: We thank Bill Kovacic, Giuliana Palumbo, Salvatore Piccolo, and Patrick Rey for many useful discussions related to this project. Spagnolo is grateful to the Wallander and Hedelius foundation for financial support (Handelsbanken P2013-0162).

♥ LEAR

♦ Università di Napoli Federico II, Department of Economics and CSEF. Email: giovanni.immordino@unina.it;

♣ SITE - Stockholm School of Economics, EIEF, Tor Vergata, CEPR. Email: spagnologianca@gmail.com

Table of contents

1 Introduction

Discovering crimes by eliciting already existing information from witnesses or accomplices may be - at least in some circumstances - more efficient than obtaining new information through costly investigations. This is why most legal regimes in history have had implicit or explicit forms of incentives for whistleblowers. However, explicit, structured financial incentives for whistleblowers, while increasingly popular in the US, are viewed with suspicion in Europe. Disagreement in the policy debate touches many aspects, but one argument always put forward by opponents of such schemes is that they will induce fraudulent reports based on false or fabricated information. It is argued that in the end, this will make law enforcement more costly and less effective. However, this does not seem to have been a major problem in the US, where the agencies administering some of these schemes argue that they are game-changers that, in addition dramatically increasing detection rates, also largely pay for their administration costs.

To shed light on this policy debate, in this paper we develop a model of the interaction between rewards for whistleblowers, sanctions for fraudulently reporting false or fabricated information, judicial errors, and standards of proof in the two court cases: the one based on a whistleblower's allegations; and the following one, typically neglected in the debate, for defamation or perjury against a whistleblower who lost the first case. Indeed, the policy debate, which has been particularly hot in recent times due to the many episodes of large-scale corporate misbehavior (from Siemens' systematic bribery to Madoff's ponzi scheme, and from Wells Fargo's accounts misselling to Volkswagen emissions cheating), overlooks that these schemes can be designed in many different ways, and that every legal system already has tools in place specifically meant to prevent fraudulent claims based on false information, like defamation and perjury laws. It is natural to start from the presumption, therefore, that the risk of an increase in false/fabricated claims can be countered by strengthening these specific tools, rather than by giving up the benefits from a potentially effective enforcement instrument like whistleblower rewards.

Our analysis shows, however, that these issues are rather subtle if one takes into account that judgement errors may occur, and that the presence and size of financial incentives and sanctions against false accusations may be reflected in changes in the court's standard of proof, and through it in the value of the information provided by the whistleblowers and in the effectiveness of enforcement.

Policy debate and institutional framework. While whistleblower reward schemes to counter federal procurement fraud and tax evasion have long been present in the US, the recent policy debate has been triggered by the financial scandals at Enron, Tyco and WorldCom, in which whistleblowers played a crucial role. This led the US Congress to enact the Sarbanes-Oxley Act (SOX) that, among other things, provided extensive whistleblower protection against retaliation. The 2007-2009 financial crisis and Wall Street bailout, followed by the uncovering of cases of large-scale financial misbehavior, like the Libor and Forex conspiracies and the systematic money laundering of blood-ridden drug cartels' revenues by HSBC, led Congress to go further and introduce within the Dodd-Frank Act provisions allowing whistleblowers to receive monetary rewards/bounties for bringing information to financial regulators.[1] Whistleblower reward programs have since been enacted by the Securities and Exchange Commission (SEC) and the Commodity Futures Trading Commission (CFTC), providing a bounty of 10-30% of the sanctions for all tips resulting in SEC36 or CFTC37 enforcement actions with monetary sanctions greater than $1,000,000. These policies are considered a success by the authorities that manage them.[2]

These programs continue to be criticized from both sides of the bar. On the one side, some legal scholars criticize the new schemes for being too weak, falling short of the "ideal" program, the False Claims Act, which has successfully provided rewards to whistleblowers that uncovered federal fraud for a long time. The claimed weakness of these new programs is that they do not allow whistleblowers to litigate cases independently of the SEC's decision, giving too much authority to institutions that could be influenced by the financial sector in a number of ways (e.g. Rapp, 2012).[3]

On the other side, some financial institutions are unhappy because these programs undermine the incentives to report using the "internal whistleblowing channels" introduced by

[1] These programs are different but related to the False Claim Act, the first reward system for whistleblowers introduced in the US, applying to federal procurement fraud. Originating from legislation enacted during the Civil War by President Lincoln to fight corruption in military procurement, the act was reinvigorated in 1986 and allows whistleblowers that provide information and win a court case on the uncovered fraud to obtain a reward up to 30% of all fines and funds recovered by the federal government. Another whistleblower rewards schemes that was already active before Dodd-Frank is that managed by the IRS to uncover tax evasion. See Engstrom (2016) for a review of the main differences across these programs.

[2] For example, in her speech at the Garret Institute in April 2015, Mary Jo White, the SEC Chairman, argued with respect to the SEC's whistleblower awards program, established by the Dodd-Frank Act in 2010, that: "The program, while clearly still developing, has proven to be a game changer." According to the SEC, many hints and successful enforcement actions have already been linked to the whistleblower reward program.

[3] For example, revolving doors. A recent high-level whistleblower from Deutsche Bank who was awarded a large reward for unveiling mismanagement rejected the reward, protesting against the inaction of the SEC towards Deutsche Bank and the fact that top officials of that bank potentially involved in mismanagement ended up at the SEC while the investigation was ongoing (Financial Times, August 18, 2016).

compliance programs encouraged by the Sarbanes-Oxley Act.[4]

While some of this criticism is understandable, the origin of the skepticism from regulatory authorities on the other side of the Atlantic is less clear. For example, in 2014 the two main UK financial market watchdogs at the time, the Bank of England's Prudential Regulation Authority and the Financial Conduct Authority, wrote a joint response (BoE-PRA and FCA, 2014) to a request for opinion from the financial market committee of the UK parliament offering widespread but apparently unsubstantiated criticism of financial incentives to whistleblowers, while neglecting the empirical evidence available (some of which published in leading scientific journals) from the US experience.[5]

This paper. In this study we focus on the concern, often raised in this debate by opponents of whistleblower rewards, that financial incentives for whistleblowers will generate a large number of "malicious" reports based on false or fabricated information for the purpose of obtaining the reward.[6] The long and successful US experience with the False Claim Act demonstrates that these potential problems can be controlled for. The interesting question is how to do this optimally. The risk that high-powered incentives lead to many

[4]These programs, though, will be relevant mostly after a corporate crime is committed, i.e. after the compliance program has failed in its task of preventing corporate crime, and are therefore at risk of being used to conceal the crime from law enforcers, protecting companies from sanctions and reducing their incentives to avoid unlawful behavior.

[5]The 2014 Note by the Bank of England and the Financial Conduct Authority argues, among other things, that in contrast to what is argued by the US agencies running these programs, whistleblowers rewards programs are bound to be ineffective, but it provides no evidence in support of this claim. More importantly, the note argues that there is no evidence to suggest that these programs may be effective (key point b, p. 2): "*There is as yet no empirical evidence of incentives leading to an increase in the number or quality of disclosures received by the regulators.*" This appears to be in contrast with the several pieces of evidence widely available at the time:

i) a very well known study by Dyck, Morse and Zingales (2010), circulating as a working paper since 2007, showing empirically that whistleblowers rewards under the False Claim Act were actually highly effective in eliciting whistleblowing from employees;

ii) the empirical studies by Engstrom (2012, 2013, 2014), published in distinguished and publicly available legal journals, providing additional empirical evidence of positive effects of the whistleblower reward system under the False Claim Act;

iii) the empirical evidence in Baloria, Marquandt and Wiedeman (2015, publicly available on SSRN.COM since 2011) showing that opposition to whistleblower rewards came mostly from companies with dubious corporate governance; and

iv) the experimental evidence in the published work of Bigoni et al. (2012) and in the work by Abbink and Wu (2017), publicly available since 2013, suggesting that only rewards are effective in inducing whistleblowers to speak out and, most importantly, in deterring misbehavior through the fear of whistleblowing.

[6]For example, Howse and Daniels report that "it is often claimed that the prospect of large awards to whistleblowers provides an incentive for employees to fabricate claims of wrongdoing for personal profit." (1995, p.540). According to a recent report by Transparency International, "Employers' representatives have been reluctant to publicly welcome the bill and have claimed it may open them to reputational damage and false or malicious claims." (2013, p.54). A report by the law firm DLA Piper (2015) on whistleblowing also stresses the risk of "malicious or unfounded allegations" against employers. Analogous concerns are expressed in the above-mentioned 2014 note by the Bank of England.

fabricated accusations can be controlled directly through more severe sanctions against defamation, perjury and information fabrication.[7] It could also be - and is expected to be controlled also by the adoption of a stricter standard of proof by courts that evaluate information coming from witnesses who stand to gain from a conviction, an effect that may offset the increased incentive to forge information linked to rewards, at the cost of a lower conviction rate.[8] Alternatively, it could be controlled by lowering or capping the rewards for whistleblowers. All these strategies entail costs in terms of weaker enforcement/reduced deterrence, so the shape of the optimal policy will likely involve a mix of strategies and is not obvious. Moreover, the efficiency/precision of the legal system is likely to play an important role, so that different policies may be optimal in different institutional environments.

As mentioned, our model focuses on the interaction between financial rewards for whistleblowers, sanctions for reporting false/fabricated information (perjury, defamation), judicial errors, and standards of proof in two court cases: one based on a whistleblower's allegations, and the potential follow-up case against a whistleblower accused of reporting false/fabricated information. We use the model first to bring to light the trade-offs involved and to derive the optimal law enforcement policy in terms of whistleblower rewards, sanctions against misreporting by a whistleblower and the standards of proof, under various parameter configurations. We then extend the model to consider the possibility of reporting through internal whistleblowing channels (part of corporate compliance programs), and how these channels could be used by guilty firms to "cover up" infringements by its managers by offering benefits ("bribes") to the employee and preventing the information on the infringement from reaching law enforcement agencies.

Our results show that to design an effective reward program, where whistleblowers do not deliberately present false claims but do file reports when a firm misbehaves, it is crucial to find an optimal balance between the reward offered to successful whistleblowers and the

[7] Some suggest that this is how this problem is avoided in the US, although additional provisions may be needed in other institutional settings. For example, Amanda Rose writes: "While it is unlikely that the WBP will produce many tips that are outright fabrications (the requirement that whistleblowers submit their tips under penalty of perjury, as well as the requirement that the SEC bring a covered action based on the tip before a bounty may be awarded, seems a sufficient guard against this), the lure of a large bounty could motivate individuals to view honest conduct as suspicious." (2014. p. 1283)

[8] The expectation of a sizable tightening of the standard of proof applied by courts when hearing testimonies from financially rewarded whistleblowers appears to be one reason why officials in the Antitrust Division of the Department of Justice (the institution that works so effectively with the False Claim Act) recently opposed a proposal to introduce a whistleblower reward program in antitrust (see, e.g. GAO 2011). This is a somewhat strange concern for the case of cartels, because these involve many firms and several individuals in each firm. Rewards could be awarded to one (innocent) informed individual who reports a cartel, only to detect new cartels. Then many other individuals, such as the leniency applicants that are not rewarded, can continue testifying in court, as it is the case today. This concern may make sense, instead, for frauds and other crimes where there are not as many other non-rewarded individuals that could be called to witness.

sanctions against those convicted for knowingly reporting false information. Given other parameters, a balanced ratio between these two parameters is the crucial tool that can lead to an optimal program, suggesting that sufficiently severe sanctions against whistleblowers convicted for forging information are necessary to compensate for large rewards.

When taking into account retaliation from employers and the fact that protection from retaliation is necessarily incomplete (not all forms of retaliation can be observed by a court), we show that there is a minimum size of reward, below which whistleblowing will not take place at all, independent of the above ratio. In these rather realistic cases, the reward must substantially exceed the sanction imposed on a mendacious whistleblower the more so the larger the risk that a firm will retaliate and the weaker the whistleblower protection.

The choice of the standard of proof is also crucial. Given the level of accuracy, an overly strict standard of proof, with very few false positives and therefore many false negatives, induces whistleblowers to never report and the firm to always commit wrongdoing. Similarly, when the standard of proof is too low, the many false positives induce whistleblowers to always blow the whistle, whether or not the firm is guilty, again leading the firm to always commit wrongdoing. The standard of proof must therefore be chosen with due care to avoid turning these schemes into failures.

Improving courts' accuracy - which reduces both types of error - is the only policy with unambiguously positive effects in terms of relaxing the constraints for a whistleblower reward scheme to be optimal. This policy is costly of course, but a well functioning whistleblower reward scheme may reduce detection and prosecution costs and at the same time increase recoveries, thereby compensating for the costs of increased accuracy. From a short-run perspective, however, these results imply that whistleblowing reward schemes are useful instruments only when the court system is sufficiently precise, and that they are likely not to perform as one would hope in environments with weak institution, where court precision is low.

As for internal whistleblower channels, we find that when the top management is involved in the wrongdoing, these channels can be easily misused to undermine external whistleblower channels. They may allow the management of guilty firms to arrange cover-ups, i.e. to 'bribe' the employee that reports internally so that the information on the firm's wrongdoing does not reach law enforcement agencies. As a consequence, the programs should not require the whistleblower to first report internally, as this would make the external whistleblower program ineffective with respect to wrongdoing by the firm management.

The rest of the paper unfolds as follows. In Section 2 we discuss the related literature. Section 3 presents the baseline model. Section 4 presents our results in the 'no retaliation' benchmark, while Section 5 does the same in the case with retaliation. Section 6 takes

into account two separate and potentially different trials: one where the firm is judged for illegal behavior, and another where the employee is judged for filing a false case. Section 7 studies the interaction between internal and external whistleblower channels. Section 8 briefly concludes.

2 Related literature

Informal discussions on the costs and benefits of rewarding whistleblowers financially abound but, despite the relevance of the issue and the lively policy debate, with extremely divided views across the Atlantic, there has been little formal economic analysis.[9] The first formal economic analyses of rewards programs are those on accomplice-witness whistleblowers. Spagnolo (2004) analyzes these programs within a dynamic model of collusion that captures the strategic features of any illegal relationship with hold-up problems within the criminal team. It is shown that offering a reward to the first self-reporting party financed by the fines paid by the remaining parties generates additional deterrence effects through "distrust" and - with finite fines - can lead to the first best of full deterrence (with zero probability of inspection from law enforcers). Aubert, Kovacic and Rey (2006) considerably extends the study of rewards to whistleblowers in collusion cases, focusing on firms internal organization and rewards for individual employees blowing the whistle on their firm's misbehavior. They emphasize the significant additional deterrence effects these schemes may bring about, arguing strongly for their introduction, but also study their possible adverse effects on firms' behavior, in particular on turnover, the incentives to innovate and cooperation, and how to minimize them.

Friebel and Guriev (2012) study rewards for innocent/bystanders whistleblowers focussing on accounting management (e.g. overstatement of financial results). They also show that besides deterring such unlawful behavior by making it more costly for the management, such rewards may also have negative effects on firms' productive efficiency by limiting their ability to give managers high powered incentives. Felli and Hortala-Vallve (2016) show how rewarding whistle-blowing can be used as a tool to prevent opportunistic behavior in the form of collusion or blackmail on the part of members of a hierarchical structure. Piccolo and Immordino (2016) study leniency programs against organized crime, and suggest that when a boss can design complex internal rules that reward his agents based on the quality of the information they can disclose, the legislator must rely on rewards to

[9] The costs and benefits of rewarding whistleblowers have been discussed informally in, for example, Howse and Daniels (1995). A highly influential article in the Washington Post by Luigi Zingales (2004) argued strongly in favor of rewarding whistleblowers who uncover financial fraud, as is currently happening at the SEC. Kovacic (2000) argued in favor of introducing these schemes in antitrust. There is an extensive legal literature on whistleblower rewards, recently surveyed in Engstrom (2016).

induce agents to report information. The most recent theoretical analysis, and the closest to our paper, is Givati (2016). This paper studies the optimal size of whistleblower rewards in a model where whistleblowers bear a personal cost, and where a reward may encourage false reports. He finds a non-monotonic relationship between the personal cost to whistleblowers and the optimal reward. A similar relationship is uncovered between the risk of a false report and the optimal reward. He also shows that when the risk of a false report is sufficiently small, whistleblowing dominates policing as a law enforcement strategy. Our model is complementary, as while he studies closely related issues, Givati (2016) does not study the possibility that a fraudulent report is discovered and sanctioned through the court system (most whistleblower schemes require disclosure information under penalty of perjury), nor how court may or should react to rewards of different size by adjusting the standard of proof, and with it the probability of different types of mistakes, for different levels of courts precision.[10]

On the empirical side

The first systematic empirical study of reward programs for whistleblowers is by Dyck, Morse and Zingales (2010). The authors investigate empirically who blows the whistle on corporate fraud by assembling and analyzing data on all reported fraud cases in large U.S. companies between 1996 and 2004. They find that corporate fraud is most often reported by employees and the media. With respect to financial incentives, they find that in sectors where whistleblower rewards are available through the False Claim Act, like the health industry, fraud is uncovered thanks to employees blowing the whistle in 41% of cases. This percentage falls to 14% in industries in which whistleblowers rewards are not available, a statistically highly significant difference.

Baloria, Marquandt and Wiedeman (2015) conduct an event study analysis showing that firms whose management lobbied against the implementation of the Dodd-Frank Act's whistleblower program, and in particular those with poorer internal compliance programs, experienced higher excess stock returns around events related to the implementation of the

[10] Somewhat less related is Hayes and Kapur (2008), examining theoretically how regulators should respond to whistleblowers' tip-offs, but assume that monetary rewards are *not* present and focus on alternative psychological theories of why they blow the whistle in the absence of financial incentives to do it.

Mechtenberg, Muehlheusser and Roider (2017) study the possibility that even whistleblower *protection* is abused by low productivity employees to postpone termination by filing fraudulent claims, suggesting that it may be efficient to let protection only start after the court has judged on the whistleblower's claim. It assumes, however, that judicial proceedings are instantaneous, so that firms cannot retaliate against whistleblowers while the court case is ongoing (many years in many countries); and that whistleblowers cannot be punished for filing fraudulent claims (although innocent firms will have a significant probability of winning in court against them, and all the resources and incentives - disciplining other employees - to pursue fraudulent whistleblowers as harshly as possible). We conjecture that in their model more realistic assumptions would lead to different results and policy prescriptions.

new rules. They also find that these new rules significantly increased firms' valuations on average. Taken together, these results indicate that investors expected the new whistleblower provisions introducing financial incentives to provide considerable net benefits by improving shareholder protection, in particular in poorly governed firms.

Wilde (2017) provides evidence that whistleblowing deters financial misreporting and tax aggressiveness. Using a dataset of retaliation complaints filed with OSHA between 2003 and 2010 on violations of paragraph 806 (which outlaws retaliation against employees who provide evidence of fraud), it finds that firms subject to whistleblower allegations in the past exhibit less financial misreporting and tax aggressiveness. The firms were also more likely to have engaged in accounting irregularities in the years previous to the allegation compared to control firms.

Finally, regarding the effectiveness of enforcement, Engstrom (2012) studies an original dataset of more than 4,000 qui tam suits filed between 1986 and 2011. He tests (and rejects) the claim that private litigation linked to the False Claim Act and whistleblower rewards is inefficiently dominated by a few, repeat, "professional" plaintiff-relators, and finds that specialized legal firms appear to play a positive role helping to unveil larger frauds than less experienced ones.[11] Call et al. (2017) analyzes OSHA claims between 2002 and 2010 and finds that whistleblowers' involvement in financial misrepresentation enforcement actions is correlated with higher monetary sanctions for the wrongdoing firm and increased jail time for culpable executives, and that when whistleblowers are involved enforcement proceedings begin quicker. It also finds that whistleblower involvement significantly increases the likelihood that criminal sanctions are imposed.

On the experimental side

The difficulties linked to the empirical analysis of illegal behavior makes this field one where laboratory experiments can be a very useful complement, with all the usual caveats regarding external validity. Indeed, there is a rich and growing literature that studies experimentally questions related to different types of whistleblowing and to the policies for dealing with them. We can follow Spagnolo (2008) and Breuer (2013) in distinguishing between studies looking at innocent (or "watchdog") whistleblowers (i.e. bystanders or employees not involved with the crime) and studies of accomplice-witness (or "traitorous") whistleblowers (i.e. those reporting information on a crime in which they took part, or were supposed to take part in, in exchange for lenient treatment from law enforcers and/or a monetary reward). Another dimension along which these strands differ is whether the focus

[11] See also Engstrom (2013) and (2014), where the same dataset is used to study how the DoJ selects and oversees claimants under the False Clam Act and the difference between jurisprudence generated by private versus public enforcement actions.

is more on whistleblowers' motivation or on the crime-deterrence effects of different legal regimes.

Studies on rewards of innocent or watchdog (or bystander) whistleblowers focus more on whistleblowers' motivation than on their effects on crime. This strand of the literature typically finds that monetary rewards are effective in increasing the number of subjects that blow the whistle. For example, Breuer (2013) studies the effects of rewards for whistleblowers in a laboratory experiment on tax evasion. He finds a strong positive effect of rewards on subjects' willingness to blow the whistle, increasing in the size of the reward, and little sign of crowding out of non-monetary motivation. He also finds that other subjects correctly anticipate this effect, so that tax evasion is significantly lower when rewards are paid to whistleblowers. Stikeleather (2016) runs a laboratory experiment framed as a corporate theft scenario and finds a significant increase in the rate of internal whistleblowing when monetary rewards are granted to employees that expose their peers' wrongdoing. More recently, Schmolke and Utikal (2016) study the frequency of whistleblowing generated by rewards, fines for not blowing the whistle, and whether and how the enforcing authority is affected by the whistleblower's report in a neutrally framed environment; while Butler, Serra and Spagnolo (2017) study the interaction between monetary rewards, the visibility of the crime reported, and public image concerns of the whistleblowers. Most relevant for us, both studies find that, controlling for other factors, monetary rewards are very effective in increasing the probability of whistleblowing.

A second strand of the experimental literature on whistleblower rewards looks at situations where one of the parties of a potential or actual illegal transaction/relationship is offered amnesty (if the illegal action occurred) and a monetary reward for blowing the whistle and reporting the illegal transaction/relationship to law enforcers. The focus of this literature is on the effects of offering a reward to whistleblowers on the number of illegal transactions that take place, i.e. on crime deterrence rather than on individual whistleblowers' motivation. Apesteguia, Dufwenberg and Selten (2007) study rewards for whistleblowers in an experiment on illegal cartel formation in the context of static Bertrand competition. Any member of a cartel may blow the whistle on its existence, in which case cartel members must pay a fine. Incentives for whistleblowing firms are varied across treatments. The authors find that whistleblower bonuses do not reduce market prices relative to the 'no incentives' treatment, although they produce the highest likelihood of whistleblowing. In a repeated game version of an analogous leniency experiment conducted by Bigoni et al. (2012), where subjects had more occasions to experiment and learn the subtleties of this rather complex strategic environment, amnesty for the first whistleblower resulted in fewer collusive agreements but higher prices in surviving ones. Rewards for the first conspiracy member blowing the whistle (funded through the fines paid by others) led instead to very

high reporting rates, thus destabilizing collusion at an increasing pace, as predicted by the theory (Spagnolo 2004). A strong effect of whistleblower rewards on the deterrence of illegal transactions is also found in a recent lab experiment on bribery by Abbink and Wu (2017). There, the possibility for one party to obtain amnesty from prosecution and a monetary reward when blowing the whistle turns out to have a strong deterrence effect on isolated illegal transactions, although the effect of repeated relationships more limited (but still positive).

Beyond economics, there is an extensive literature on whistleblowing in the areas of law, sociology, psychology, business, and public administration that we cannot cover here (see surveys in Engstrom 2016; Miceli and Near 1992; and Miceli, Near and Dworkin 2008).

3 The baseline model

Consider a simple game between a firm which chooses whether to act legally or illegally and an employee who, after observing the firm's choice, can decide to file a case against the firm or to remain silent.

Undetected illegal behavior yields a random monetary return π, distributed on the support $(0, +\infty)$ with cumulative distribution function $H(\pi)$ and probability density function $h(\pi)$. If the employee blows the whistle (either against a guilty firm or against an innocent firm), a trial takes place without loss of generality with probability 1 and it may end up with a right or a wrong verdict. Let 0 be the normalized profits of the honest firm that is not convicted and the payoff of the employee when he does not file a case, F the monetary sanction faced by a convicted firm, R the reward the employee gets in case of a successful filing, and f the sanction imposed on an employee who is convicted of fabricating information to file a complaint. For the time being, we assume that the firm does not retaliate against whistleblowers (we will relax this assumption in Section 5).

Two errors may occur at the end of the trial: a false positive and a false negative. We indicate with α_p the probability of wrongful conviction of an innocent firm (false positive) induced by the standard of proof prescribed to courts for this situation by the current legal system; and with α_n the probability of wrongful acquittal of a guilty firm (false negative) under the same system.

We shall also assume $\alpha_n^{'}(\alpha_p) < 0$, $\alpha_n^{''}(\alpha_p) > 0$ and $\alpha_i \in [0, \bar{\alpha}]$ with $\bar{\alpha} < \frac{1}{2}$ and $i = p, n$. Note that setting the probability of conviction of an innocent firm α_p automatically determines the probability of acquittal of a guilty firm $\alpha_n(\alpha_p)$. The reason is that it is not possible to decrease one type of error without increasing the other. For details of the underlying derivation from the choice of the burden of proof and the minimum strength of evidence required to apply the sanction, see the contributions by Kaplow (2011a, 2011b,

2012).

We assume that the employee gets the reward R if the firm is convicted, and the sanction f if the firm is acquitted. Therefore α_p also represents the probability that the reward is paid to an employee who files a false claim, and $\alpha_n(\alpha_p)$ represents the probability that the employee is sanctioned when he files against a guilty firm.

The timing of the baseline game is as follows:

t =1 The return from undetected illegal behavior π materializes. The firm decides whether to commit the illegal act or not. Once the illegal act is committed, the game proceeds to the next stage.

t =2 The employee, knowing if an illegal act has been committed or not, decides whether to file a claim (false or otherwise).

t =3 The trial uncertainty resolves and sanctions are imposed.

The solution concept is subgame perfect Nash equilibrium.

All players are risk neutral, so that sanctions can be interpreted as the monetary equivalent of the imprisonment terms, fines, damages, and so forth, to which the criminals expose themselves. We also assume the following tie-breaking condition.

A1 When the employee is indifferent between blowing the whistle and remaining loyal to the firm, he chooses the latter option.

Given these variables and parameters, a whistleblower program is described by a quadruple (R, f, F, α_p). In what follows, we will comment on the design of an optimal whistleblower program.

4 The 'no retaliation' benchmark

As a benchmark, consider first the case in which the firm is not allowed to retaliate against whistleblowers.

In order to solve the model, suppose that (along the equilibrium path) the firm has committed an illegal act. Then, the employee will file a case if and only if

$$u_I = (1 - \alpha_n(\alpha_p))R - \alpha_n(\alpha_p)f > 0 \Leftrightarrow \frac{(1 - \alpha_n(\alpha_p))}{\alpha_n(\alpha_p)} > \frac{f}{R}.$$

If, instead the firm has not committed an illegal act then the employee will file a false claim if and only if

$$u_L = \alpha_p R - (1-\alpha_p)f > 0 \Leftrightarrow \frac{\alpha_p}{1-\alpha_p} > \frac{f}{R},$$

where u_I and u_L are the employee's expected utility in each subgame. Note that

$$S_I \equiv \frac{(1-\alpha_n(\alpha_p))}{\alpha_n(\alpha_p)} \text{ and } S_L \equiv \frac{\alpha_p}{1-\alpha_p}$$

are the odds of success, defined as the ratio of the probability of success over the probability of failure, from filing a case with and without merit, respectively. Moreover, $S_I > 1 > S_L$ because $\alpha_i < \frac{1}{2}$. Then, depending on the design of the whistleblower program (R, f, F, α_p) there are only three possible cases. In the first case, the employee always blows the whistle (this is the case if $S_I > S_L > f/R$). In the second case the employee blows the whistle only when the firm misbehaved (i.e. $S_I > f/R \geqslant S_L$. Finally, the employee never blows the whistle if $f/R \geqslant S_I > S_L$).

Going backward, we now characterize the firm's decision to commit an illegal act. The firm's profit from committing an illegal act is π if the employee does not blow the whistle and $\pi - F(1-\alpha_n(\alpha_p))$ otherwise. Instead, profits from behaving legally are 0 if the employee does not blow the whistle and $-F\alpha_p$ otherwise. Next, note that if $S_I > S_L > f/R$, the employee always blows the whistle, and the firm decides to act illegally or not by comparing $-F\alpha_p$ and $\pi - F(1-\alpha_n(\alpha_p))$. In this case, the crime is committed if and only if the return exceeds the following threshold

$$\underline{\pi} \equiv F(1-\alpha_n(\alpha_p)-\alpha_p),$$

which is increasing in the firm's sanction F and in the probability of conviction of an innocent firm (α_p) whenever increasing false positives decreases false negatives more than proportionally $(\alpha_n'(\alpha_p) < -1)$. If instead $S_I > f/R \geqslant S_L$, the employee blows the whistle only when the firm misbehaves and the firm decides to act illegally or not by comparing 0 and $\pi - F(1-\alpha_n(\alpha_p))$. In this second case, the crime is committed if and only if the return exceeds the following larger threshold

$$\bar{\pi} \equiv F(1-\alpha_n(\alpha_p)),$$

which is increasing in the firm's sanction F and in the probability of conviction of an innocent firm α_p. Finally, if the employee never blows the whistle $(f/R \geqslant S_I > S_L)$, the firm will always commit the illegal act since $\pi > 0$. Putting together our findings above, we get a complete characterization of the (subgame perfect) equilibrium of our model.

Proposition 1 *If $S_I > S_L > f/R$, the employee always blows the whistle and the crime rate is*

$$\Pr[\pi \geqslant \underline{\pi}] = 1 - H(F(1 - \alpha_n(\alpha_p) - \alpha_p)),$$

which is decreasing in the firm's sanction F and in the probability of conviction of an innocent firm α_p whenever $\alpha_n^{'}(\alpha_p) < -1$.

If $S_I > f/R \geqslant S_L$, the employee blows the whistle only when the firm misbehaves and the crime rate is

$$\Pr[\pi \geqslant \bar{\pi}] = 1 - H(F(1 - \alpha_n(\alpha_p))),$$

which is decreasing in the firm's sanction F and in the probability of conviction of an innocent firm α_p.[12]

Finally, if $f/R \geqslant S_I > S_L$, the employee never blows the whistle and the crime rate is 1.

The area where the probability of successfully filing a case is larger than f/R, both following legal and illegal behavior by the firm, represents whistleblower programs that induce an employee to always file a case, including against honest firms, in which case the firm optimally decides to misbehave or not depending on the sanction and on the errors which may occur at the end of the trial. The area where the probability of successfully filing a case is larger than f/R following illegal behavior, but is smaller for legal behavior represents whistleblower programs where employees denounce guilty firms and the firm behaves legally if the firm's sanction F and the probability of conviction of an innocent firm α_p are sufficiently high. Finally, if the probability of successfully filing a case is always smaller than f/R, the program never provides an incentive to file a case, so that the firm always acts illegally.

From the point of view of society, a whistleblower program is optimal if it induces an equilibrium in which the firm acts legally and employee only files case against a guilty firm. Concerning the employee's behavior, it is then optimal to choose (R, f, F, α_p) so as to induce the second case. A few comments are in order.

First, it is clear that rewards for honest whistleblowers and sanctions against dishonest ones must be strictly related. The danger from fixing the wrong combination of these two policy parameters is disrupting the entire program. In the case that the incentive structure of the program is such that the employee always acts against the firm, and even more so

[12] Since, for simplicity, we are assuming that firms will only be prosecuted if a whistleblower reports incriminating information/files a claim against them, in the specific cases we will of course have $\alpha_p^i = 0$ when in situation i the whistleblowers never file claims against innocent firms, and $\alpha_n^j = 1$ when in situation j whistleblowers never file claims against a guilty firm either. The comparative statics on the error probabilities is possible because it is on the legal system's parameters $\alpha_p(\alpha_n)$, not on the probability in each specific case.

in the opposite case, firms may as well act illegally, counting on a defective policing and judiciary system.

Second, for very low value of f, the range of rewards, R, that allow the first best to be obtained is very narrow, but it gets larger and larger as long as f increases. This implies that monetary sanctions against mendacious whistleblowers may not be sufficient, as they are bound to be limited to their personal wealth and may prove too lenient, especially when high rewards for successful cases are envisaged. In addition, a corollary of proposition 1 is that a sufficient condition to have an optimal whistleblower program, as far as the two parameters R and f are concerned, is to set $R = f$. If the reward, R, is a percentage of the firm's illicit gain, as it is in some actual programs, then it can easily become substantial and it is likely that a monetary sanction for a mendacious whistleblower, which is constrained by his personal wealth, is insufficient to provide incentives that yield the desired behavior. It follows that the bounty program must provide for criminal sanctions against a whistleblower who files a fraudulent claim.

Third, the choice of the burden of proof and the minimum strength of evidence required to apply the sanction is extremely important to make the program work. To see this, consider that both S_I and S_L decrease as the probability of conviction of an innocent firm α_p decreases. In other words, an imperfect setting of the burden of proof could change the program from one where employees always blow the whistle to one where they never do (or vice versa), missing the first-best area.

Even if whistleblower programs are seen as a mechanism to elicit information that would not be available otherwise, courts must not rely only on information provided by relators. Whistleblowers programs have to be seen as complements to other interventions that improve the accuracy of the judiciary system. As explained by Kaplow (2011b), accuracy determines the overall error rate of the legal system, whereas the burden of proof dictates how to divide the errors between the two types. This also shows that a whistleblower program may prove very dangerous if it is introduced in countries where, due to corruption or to an inefficient system, the performance of courts is unsatisfactory. We now explain by way of a thought experiment why accuracy is another important dimension in the design of the judiciary system. Assume that $S_I > S_L > f/R$ and that the f/R cannot be modified. Then we are in the area where the employee always blows the whistle. Suppose that we can only change the burden of proof to switch to the first-best scenario where $S_I > f/R \geqslant S_L$ so that the employee blows the whistle only when the firm misbehaves. If we decrease the probability of conviction of an innocent firm α_p, then S_L will decrease going in the right direction. However, decreasing the probability of a false positive will increase the possibility of a false negative, so that S_I will also decrease. Therefore, depending on the relative speed at which S_L and S_I decrease, we might enter the worst-case scenario, i.e. $f/R \geqslant S_I > S_L$,

where the employee never blows the whistle and the crime rate is maximal. However, this worst-case scenario can be easily avoided if in improving accuracy we can simultaneously decrease α_p and $\alpha_n(\alpha_p)$, so that the change in the burden of proof is counterbalanced by an overall improvement in the judicial system.

5 The case of retaliation

Whistleblowers are subject to various forms of retaliation, many of which cannot be covered by whistleblower protection laws where these are present (for instance, slowing down career progression, through delayed promotions, or blacklisting/lack of offers from other firms in the industry). It is therefore of first-order importance to study the impact of firm retaliation on the simple result highlighted in the previous section.

In the previous section, the reward R was implicitly assumed to be net of any cost. In practice, the main cost for the employee is the risk of reprisal by the firm. In this section, we address this issue explicitly by introducing the possibility of retaliation measured by a parameter P, the punishment a whistleblowers is subject to when the reported entity retaliates.[13]

Going through the same steps as in the previous section, we find two new conditions such that the employee will file a case when the firm has committed an illegal act

$$u_I^P = (1-\alpha_n(\alpha_p))R - \alpha_n(\alpha_p)f - P > 0 \Leftrightarrow S_I \equiv \frac{1-\alpha_n(\alpha_p)}{\alpha_n(\alpha_p)} > G,$$

where $G \equiv \frac{f+P}{R-P}$, and when the firm has not committed an illegal act

$$u_L^P = \alpha_p R - (1-\alpha_p)f - P > 0 \Leftrightarrow S_L \equiv \frac{\alpha_p}{1-\alpha_p} > G,$$

where u_I^P and u_L^P are the employee's expected utility in each subgame. Then, depending on the design of the whistleblower program (R, f, F, α_p, P), there are only three possible cases. In the first case, the employee always blows the whistle, this is the case if $S_I > S_L > G$. In the second case, the employee blows the whistle only when the firm misbehaves, i.e. $S_I > G \geqslant S_L$. Finally, the employee never blows the whistle if $G \geqslant S_I > S_L$.

Going backward, the firm's decision to commit the illegal act is exactly as before for each of the three areas described above.

Putting together our findings, we obtain a complete characterization of the (subgame perfect) equilibrium of our model.

[13] This parameter can also be interpreted as the inverse of the degree of whistleblower protection allowed by a legal system, with $P = 0$ being the (unlikely) case of perfect protection from retaliation.

Proposition 2 *If $S_I > S_L > G$ the employee always blows the whistle and the crime rate is*

$$\Pr[\pi \geqslant \underline{\pi}] = 1 - H(F(1 - \alpha_n(\alpha_p) - \alpha_p)),$$

which is decreasing in the firm's sanction F and in the the probability of conviction of an innocent firm α_p whenever $\alpha_n^{'}(\alpha_p) < -1$.

If $S_I > G \geqslant S_L$, the employee blows the whistle only when the firm misbehaves and the crime rate is

$$\Pr[\pi \geqslant \bar{\pi}] = 1 - H(F(1 - \alpha_n(\alpha_p))),$$

which is decreasing in in the firm's sanction F and in the probability of conviction of an innocent firm α_p.

Finally, if $G \geqslant S_I > S_L$ the employee never blows the whistle and the crime rate is 1.

Although Propositions 1 and 2 are very similar, there is one important difference. The clear and important message of the previous result was that rewards for honest whistleblowers and sanctions against dishonest ones must be strictly related. The absolute value of rewards for filing a merit case R and the sanction for filing one without merit f were not important, but the ratio between the two was very important. This result was due to the lack of any personal costs from blowing the whistle for the employee. More realistically, we now observe that there will never be reporting if the reward R is not sufficiently larger than the expected punishment P, since otherwise the employee's expected utility from filing a case will always be negative. Indeed, again a corollary of proposition 2 is that the optimal behavior of the potential whistleblower is induced when the program envisages a combination of reward and sanction such that $G = 1$. However, in the new setting this entails $R = f + 2P$, so that the reward must substantially exceed the sanction imposed on a mendacious whistleblower, the more so the larger is the risk that a firm will retaliate against the relator whenever turned in and whenever the protection provided to whistleblowers is weak.

6 Criminal versus administrative trial

Until now, we have kept things simple by assuming that there is only one trial where either the firm is convicted for illegal behavior, or the employee is convicted for filing a false case (e.g. fabricating information). In reality, of course, there would be two separate processes, and there is the possibility that neither the firm nor the employee is convicted. Moreover, while the first trial may be a purely administrative one, leading to a fine against the corporation, the second one will certainly be a criminal trial, as it involves perjury/lying to the court. Complicating the model by taking into account the two separate and potentially

different trials does not dramatically affect our previous conclusions, but it allows us to study how different combinations of efficiency/precision of the legal system for administrative and criminal offenses affect the performance of whistleblower schemes.

As before, let 0 be the profits of the honest firm that is not convicted and the payoff of the employee when he does not file a case, π the profits of undetected illegal behavior, F the monetary sanction faced by a convicted firm, and R the reward the employee gets in case of a successful filing, net of any personal costs. Again, two errors may occur at the end of a trial: a false positive and a false negative. We continue to indicate with α_p the probability of conviction of an innocent firm (false positive) and with $\alpha_n(\alpha_p)$ the probability of acquittal of a guilty firm (false negative). We assume that the employee gets the reward R if the firm is convicted. If the firm is acquitted, a second trial takes place.

This second trial is against the whistleblower, accused of fabricating false information and, as with the first trial, may lead to two types of error: a false positive with probability β_p, the probability that an innocent whistleblower is convicted; and false negative with probability $\beta_n(\beta_p)$, the probability of acquittal of a whistleblower who filed a false claim fabricating information. If the employee is convicted in this second trial, he faces a sanction denoted with f. Also for the second trial we assume that $\beta_n^{'}(\beta_p) < 0$, $\beta_n^{''}(\beta_p) > 0$ and $\beta_i \in [0, \bar{\beta}]$ with $\bar{\beta} < \frac{1}{2}$ and $i = p, n$ Moreover, we assume that α_p and β_p are independent.

We consider the more general case in which the firm can retaliate against the employee imposing a harm P. The case with no retaliation can be obtained by simply setting $P = 0$. As mentioned before, the parameter P can also be interpreted as the (inverse of) the protection that the program is able to guarantee to employees that are willing to cooperate with the enforcers. We can follow the same steps as in the two previous sections to identify the conditions that would induce the employee to file a case. If the firm acts illegally, the employee blows the whistle if

$$u_{2I} = (1 - \alpha_n(\alpha_p))R - \alpha_n(\alpha_p)\beta_p f - P > 0 \Leftrightarrow S_I \equiv \frac{(1 - \alpha_n(\alpha_p))}{\alpha_n(\alpha_p)} > \frac{\beta_p f + P}{R - P} \equiv G_I,$$

whereas he files a case, when the firm acts legally if

$$u_{2L} = \alpha_p R - (1 - \alpha_p)(1 - \beta_n(\beta_p))f - P > 0 \Leftrightarrow S_L \equiv \frac{\alpha_p}{1 - \alpha_p} > \frac{(1 - \beta_n(\beta_p))f + P}{R - P} \equiv G_L \ ,$$

where u_{2I} and u_{2L} are the employee's expected utility in each subgame that now includes two trials. We have to note that, given the assumptions on β_i, $G_L > G_I$. Hence, again there are three possible cases. First, the employee always files a case if $S_L > G_L$. Note that this condition implies that $S_I > G_I$. In the second case, the employee never files a case; this

occurs if $S_I \leqslant G_I$, which implies $S_L < G_L$. Finally, the employee blows the whistle only against a firm that acted illegally if $S_I > G_I$ and $S_L \leqslant G_L$.

These conditions are summarized in the following proposition.

Proposition 3 *The employee always blows the whistle if $S_L > G_L$ and the crime rate is*

$$\Pr[\pi \geqslant \underline{\pi}] = 1 - H(F(1 - \alpha_n(\alpha_p) - \alpha_p)),$$

which is decreasing in the firm's sanction F and in the probability of conviction of an innocent firm α_p whenever $\alpha_n^{'}(\alpha_p) < -1$.

The employee blows the whistle only when the firm misbehaves if $S_I > G_I$ and $S_L \leqslant G_L$ and the crime rate is

$$\Pr[\pi \geqslant \bar{\pi}] = 1 - H(F(1 - \alpha_n(\alpha_p))),$$

which is decreasing in the firm's sanction F and in the probability of conviction of an innocent firm α_p.

Finally, the employee never blows the whistle and the crime rate is 1if $S_I \leqslant G_I$.

Of course the last proposition is the most general one, as all the others can be seen as the special cases that occur when some specific conditions hold. Hence, we can make some final remarks on the basis of the results described in this proposition. First, we have to note that the simple sufficient conditions for a program that induces the employee to blow the whistle only against guilty firms are no longer valid. However, in this more general case we can exploit the fact that $S_I > 1 > S_L$ to identify a similar sufficient condition. Indeed, the whistleblower program induces an optimal behavior by the employee if $G_I \leqslant 1$ and $G_L \geqslant 1$. Solving the first inequality, we obtain $R \geqslant 2P + \beta_p f$, whereas the second inequality is satisfied if and only if $R \leqslant 2P + (1 - \beta_n(\beta_p))f$. Hence, the sufficient condition to obtain the desired behavior from the employee is:

$$2P + (1 - \beta_n(\beta_p))f \geqslant R \geqslant 2P + \beta_p f \ . \tag{1}$$

This condition would also guarantee that some policy instruments work in the expected direction, as we will see in a moment.

A whistleblower program can be described by the following policy parameters $(R, f, F, \alpha_p\beta_p, P)$ which identify the available policy instruments: 1) the reward offered to the whistleblower that helps uncover an illegal corporate act; 2) the sanction imposed on the whistleblower who fabricates information; 3) the sanction imposed on the firm that acts illegally; 4) the accuracy of the judiciary in deciding on the allegations made against either the firm or the whistleblower; 4) the standard of proof that is required in the two trials to

convict either the firm or the whistleblower; 5) the level of protection granted to whistleblowers against possible retaliations by firms. The optimal combination of these policy instruments depends on the available choice sets, their relative costs and their effectiveness and cannot be established without adding considerable structure to the model. However, we can describe how the various policy instruments affect the likely outcome of the program.

Let us start with their impact on S_I and S_L. These odds depend only on α_p and α_n. Therefore, the only relevant policy instruments are the accuracy of the legal system and the standard of proof. Investments made to improve the accuracy of the system will decrease both α_p and α_n. This will make S_I and S_L move apart (S_I increases and S_L decreases), increasing the probability that the whistleblower program induces an optimal behavior by the employee. In contrast increasing the standards of proof that has to be met to convict a firm will decrease α_p and increase α_n. This choice would determine a lower value for both S_I and S_L, making it less likely that the employee files a case.

Moving to G_I and G_L, we note that all the policy instruments - with the exception of the sanction against the firm - are relevant. Their impact on the effectiveness of the program is described in the following table.

Policy choice	**Impact on G_I and G_L**	**Conditions**
Increase R	Both decrease	always
Increase f	Both increase	always
Improve accuracy (decrease both β_p and β_n)	G_I decreases G_L increases	$R > P$
Increase standard of proof (decrease β_p and increase β_n)	Both decrease	always
Improve protection (decrease P)	Both increase	always

These results show that, in general, the various policy instruments must be handled with care as they might lead to a policy that is either too generous to the potential whistleblower or not generous enough. The only policy instrument that has an unambiguous impact on the incentives of the whistleblower is improving the accuracy of the judiciary, which makes it more likely that the employee blows the whistle only when the firm misbehaves.

7 Whistleblower rewards and compliance programs

In the previous models, we have shown that a whistleblower program, described by the parameters $(R, f, F, \alpha_p\beta_p, P)$, induces a socially desirable behavior if the policy instruments

represented by R (the reward to the whistleblower), f (the sanction against false claims), and P (the level of protection granted to whistleblowers) are properly combined, given the applicable standards of proof and the accuracy of the judiciary (described by α_p, α_n, β_p and β_n). Hence, it seems that the level of the sanction imposed on the firm that commits the crime, F, does not enter into the optimal design of the program. In this section, we consider an extension that shows that this may not be true.

Faced with the prospect of any fraud or misconduct issues becoming a law enforcement case —possibly even before the organization itself learns of the problem — most companies have launched compliance programs that include internal reporting systems for employees. The dominant narrative for these compliance systems is one of dishonest employees and the tool honest top management can use to control them. However, as we know from many recent scandals, from Enron to Siemens and from HSBC and Wells Fargo to Volksvagen, the primary source of misbehavior is actually top management. In these common situations, reporting through the internal compliance system may allow guilty executives to discover that a claim might be filed to the law enforcement authority regarding their misbehavior before it actually occurs. In this section, we modify the baseline model to study the interaction between an external whistleblower rewards scheme and an internal compliance program when top management is the wrongdoer and lower rank employees the potential whistleblowers. In such an extension, we assume that the firms' managers may induce the potential whistleblower not to report the alleged crime to the law enforcer by providing him with some private benefits, e.g. in the form of a career promotion, or a simple bribe. For simplicity, we refer to the private benefit the employee obtains when he accepts the firm's offer as a "bribe", as in Aubert et al. (2006).

The timing of the modified game is as follows:

t =1 The return from undetected illegal behavior π materializes. The firm decides whether to commit the illegal act or not. Once the illegal act is committed, the game proceeds to the next stage.

t =2 The employee, knowing if an illegal act has been committed or not, can either i) go through the internal compliance program and try to get a bribe for not reporting the illegal act or for not making a false claim the firm; ii) adhere to the whistleblower program and file a claim (false or otherwise); or iii) decide not to file a claim or to look for a bribe.

t =3 The trial uncertainty resolves and sanctions are imposed.

For simplicity, and with no loss of intuition, we consider the baseline case in which the firm does not retaliate against the employee, setting $P = 0$, and in which there is a single

trial.

Consider first the case where (along the equilibrium path) the firm has committed an illegal act. Then, filing a claim the employee would get u_I, as in the baseline model. Instead, if he chooses to go through the internal compliance program, he would bear a transaction cost k and would obtain a bribe (to be determined soon) B_I. Finally, the payoff from not blowing the whistle nor adhering to the compliance program is normalized to zero.

Equating the employee's payoffs from filing a claim and from blackmailing the firm, we get the minimum bribe that the employee would accept for not filing the case, i.e.,

$$\underline{B}_I \equiv u_I + k.$$

The maximum bribe that the firm will be ready to pay is instead obtained by equating the payoff which would result for a guilty firm if the employee files a claim, i.e., $\pi - F(1-\alpha_n(\alpha_p))$ and the payoff which would result from bribing the employee not to blow the whistle, i.e., $\pi - B$,

$$\bar{B}_I \equiv F(1-\alpha_n(\alpha_p)).$$

Hence, the existence of a compliance program with an internal reporting mechanism may disrupt the whistleblower program unless the following condition holds:

$$\underline{B}_I > \bar{B}_I \Longleftrightarrow R(1-\alpha_n(\alpha_p)) - \alpha_n(\alpha_p)f + k > F(1-\alpha_n(\alpha_p)). \tag{2}$$

Going through the same steps for the case where the firm has not committed an illegal act, we recover the minimum bribe that the employee would accept for not filing the false case, the maximum bribe that the firm will be ready to pay to avoid the filing, and the condition that must hold to avoid collusion between employer and employee:

$$\underline{B}_L \equiv \qquad u_L + k, \qquad \bar{B}_L \equiv F\alpha_p, \tag{3}$$

$$\underline{B}_L > \bar{B}_L \Longleftrightarrow R\alpha_p - (1-\alpha_p)f + k > F\alpha_p. \tag{4}$$

In section 4, we proved that a sufficient condition in the baseline model for having a program that determines optimal whistleblower behavior was $R = f$. So, let us assume that this equality holds. Conditions (2) and (4) become:

$$R > \frac{S_I}{S_I - 1}F - \frac{k}{\alpha_n(\alpha_p)(S_I - 1)} \tag{5}$$

and

$$R > \frac{S_L}{S_L - 1} F - \frac{k}{\alpha_p (S_L - 1)} \, . \tag{6}$$

This shows that unless k is sufficiently high, the reward that has to be paid to the whistleblower, R, has to exceed the fine imposed on the firm, F. In particular if $k = 0$, since $S_L/(S_L - 1) > S_I/(S_I - 1)$ we may have three cases: 1) $R < S_I/(S_I - 1)F$; 2) $S_I/(S_I - 1)F < R < S_L/(S_L - 1)F$; and 3) $R > S_L/(S_L - 1)F$. In case 1) the employee always reports internally a misconduct in order to obtain the bribe; in case 2) the employee only blackmails an innocent firm; and in case 3) the internal reporting system envisaged by the compliance program does not alter the effectiveness of the external whistleblower legislation. Unlike from before, now the firm's sanction F plays a crucial role in determining if the employee will file a case or not.

This extended version of the model allows an important policy question to be addressed: Should employees be required to initially report internally to their employer in order to qualify for a bounty? Such a requirement for internal reporting is not a feature of Dodd-Frank schemes, but it has been put forward as a way to make internal and external whistleblowing channels compatible. Corporations argue that internal reporting procedures would be rendered ineffective unless awards were contingent on prior internal reporting. On the other side, the whistleblowers "lobby" maintains that a requirement for internal reporting would have a chilling effect on the rewards program.

Our findings indicate that the legislator should avoid decreasing the cost k for the employee to make use of an internal compliance program. In other words, our model suggests that the lack of confidence in corporate internal controls — inherent in the Dodd-Frank legislation — is well placed. We should avoid making internal reporting easier, and employees should not be required to initially report internally to their employer in order to qualify for a bounty. Indeed, if k is sufficiently high, we are back to the baseline model with employees relying on the whistleblower program according to proposition 1. If, instead, k is low, the employee will always try to get a bribe from the firm.

Note that for some values of k it is possible that the employee might try to get a bribe in one case, and report the external whistleblower program in another. Indeed, since $(1 - \alpha_n(\alpha_p))$ is larger than α_p while $(1 - \alpha_p)$ is larger than $\alpha_n(\alpha_p)$, the employee might choose to file a true claim and blackmail the firm with a false accusation or vice versa, depending on the way the whistleblower program is designed.

8 Conclusion

Financial incentives for corporate whistleblowers are at the center of a heated debate, as US enforcement agencies have been increasingly relying on them, while their European counterparts appear weary of introducing them. One argument that is often raised by opponents of these incentives is that they will lead to an increase in fraudulent reports based on false or fabricated information that, in the end, will reduce the effectiveness of enforcement. Fraudulent claims, however, do not seem to have been a major problem for at least some of the US agencies. One plausible explanation, somewhat ignored in the policy debate, is that in a sufficiently effective legal system, fraudulent claims based on false or fabricated information have a sizable probability of failing, possibly also because the standard of proof is raised when a whistleblower expects a reward; and the whistleblowers that brought the claims can then be taken to trial for perjury or defamation. To shed light on this hypothesis, we present a model of the interaction between rewards for whistleblowers, sanctions for maliciously reporting fabricated information, judicial errors, and standards of proof in two court cases - one based on a whistleblower's allegations, and the subsequent one against the whistleblower for possibly false accusations. We use the model to analyze the optimal law enforcement policy under various parameter configurations and efficiency levels of the legal system. Our results show that an appropriate balance between the reward offered to successful whistleblowers, the sanctions against whistleblowers convicted for fabricating information, and the courts' standards of proof, is essential for whistleblower policies to succeed. When whistleblowers are financially constrained, heavy criminal sanctions for false reporting are required. The relatively milder sanctions common to European legal systems may therefore explain the opposition these schemes encounter in Europe.

When the risk of retaliation against whistleblowers is severe, higher rewards are needed for the program to be effective, and even more severe sanctions against dishonest whistleblowers become necessary. And the precision of the legal system must be sufficiently high for these programs to be viable at all. This suggests that these schemes should not be introduced in weak institutional environments, were protection from retaliation is typically less effective and the precision of the legal system is lower.

We also show that internal reporting systems that form part of firms' compliance programs may actually interfere in a rather unexpected way with the external reporting systems. Subject to agencies disclosing their data, empirical work is highly welcome and needed to test these results and appropriately shape enforcement against the increasing levels of illegality plaguing international financial transactions.

References

[1] Abbink, K. and K. Wu, (2017), "Reward self-reporting to deter corruption: An experiment on mitigating collusive bribery," *Journal of Economic Behavior & Organization,* 133, 256-272.

[2] Apesteguia, J., Dufwenberg, M. and R. Selten, (2007), "Blowing the Whistle," *Economic Theory*, 31(1),143-166.

[3] Aubert, C., Kovacic, W., and Rey P. (2006), "The Impact of Leniency and Whistleblowing Programs on Cartels," *International Journal of Industrial Organization*, 24, 1241-1266.

[4] Baccara, M. and H. Bar-Isaac, (2008), "How to Organize Crime," *Review of Economic Studies*, 75, 1039-1067.

[5] Baloria, V., Marquardt, C. and C. Wiedman, (2015), "A Lobbying Approach to Evaluating the Whistleblower Provisions of the Dodd-Frank Reform Act of 2010", unpublished working paper. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1923310.

[6] Bank of England-Prudential Regulations Authority & Financial Conduct Authority (BoE-PRA and FCA), (2014), "Financial Incentives for Whistleblowers", available for download in September 2016 until (at least) June 2017 at https://www.fca.org.uk/news/financial-incentives-for-whistleblowers.

[7] Becker, G.S., (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy,* 76, 169-217.

[8] Bigoni, M., Le Coq, C., Fridolfsson, S. and G. Spagnolo, (2012), "Fines, Leniency and Rewards in Antitrust," *RAND Journal of Economics*, 43(2), 368-390.

[9] Breuer, L., (2013), "Tax Compliance and Whistleblowing – The Role of Incentives," *The Bonn Journal of Economics,* Vol 2, No. 2, pp. 7-44.

[10] Butler, J., Serra, D. and G. Spagnolo, (2017), "Motivating Whistleblowers," Working Paper, Southern Methodist University.

[11] Call, A.C., Martin, G.S., Sharp, N.Y. and J.H. Wilde, (2017), "Whistleblowers and Outcomes of Financial Misrepresentation Enforcement Actions," Working Paper available at SSRN: https://ssrn.com/abstract=2506418, forthcoming in the *Journal of Accounting Research.*

[12] DLA Piper (2015), "Wistleblowing laws: An employer's guide to global compliance" second edition. Available at https://www.dlapiper.com/~/media/Files/Insights/Publications/2015/06/ Whistleblowing_Law_Report_2015.pdf.

[13] Dyck, A., Morse, A. and L. Zingales, (2010), "Who Blows the Whistle on Corporate Fraud?" *The Journal of Finance*, 65, 2213-53.

[14] Dworkin, T.M. and J.P. Near (1997), "Better Statutory Approach to Whistleblowing," *Business Ethics Quarterly*, 7, 1-16.

[15] Engstrom, D.F., (2012), "Harnessing the Private Attorney General: Evidence from Qui Tam Litigation," *Columbia Law Review*, 112, 1244-1325.

[16] Engstrom, D.F. (2013), "Public Regulation of Private Enforcement: Empirical Analysis of DOJ Oversight of Qui Tam Litigation Under the False Claims Act," *Northwestern University Law Review*, 107, 1689-1756.

[17] Engstrom, D.F., (2014), "Private Enforcement's Pathways: Lessons from Qui Tam Litigation," *Columbia Law Review*, 115, 1913-2006.

[18] Engstrom, D.F., (2016), "Bounty Regimes," in *Research Handbook on Corporate Criminal Enforcement and Financial Misdealing*, J.Arlen ed., Edward Elgar Press.

[19] Felli, L. and R. Hortala-Vallve, (2016), "Collusion, Blackmail and Whistle-Blowing," *Quarterly Journal of Political Science*, 11(3), 279-312.

[20] Friebel, G. and S. Guriev, (2012), "Whistle-Blowing and Incentives in Firms," *Journal of Economics & Management Strategy*, 21(4), 1007-1027.

[21] GAO-11-619 (2011). U.S. Gov't Accountability Office Report 11-619, *Criminal Cartel Enforcement: Stakeholder Views on Impact of 2004 Antitrust Reform Are Mixed, but Support Whistleblower Protection,* available at http://www.gao.gov/new.items/d11619.pdf.

[22] Givati, Y., (2016), "A Theory of Whistleblower Rewards," *Journal of Legal Studies*, 45, 43-72.

[23] Howse, R. & R.J. Daniels, (1995), "Rewarding Whistleblowers: The Costs and Benefits of an Incentive-Based Compliance Strategy," *Corporate Decision-Making in Canada,* 525-549.

[24] Heyes, A. and S. Kapur, (2009), "An Economic Model of Whistle-Blower Policy," *Journal of Law, Economics, & Organization,* 25(1), 157-182.

[25] Kaplow, L., (2011a), "On the Optimal Burden of Proof," *Journal of Political Economy*, 119, 1104-1140.

[26] Kaplow, L., (2011b), "Optimal Proof Burdens, Deterrence, and the Chilling of Desirable Behavior," *American Economic Review: Papers and Proceedings*, 101, 277-280.

[27] Kaplow, L., (2012), "Burden of Proof," *Yale Law Journal,* 121, 738-859.

[28] Kovacic, W.E., (2000), "Private monitoring and antitrust enforcement: Paying informants to reveal cartels," *George Washington Law Review,* 69, 776.

[29] Mechtenberg, L., Muehlheusser, G. and A. Roider, (2017), "Whistle-blower protection: Theory and experimental evidence", CEPR Discussion Paper No. 11898.

[30] Miceli, M.P. and J.P. Near, (1992), *Blowing the Whistle: The Organizational and Legal Implications for Companies and Employees*, Lexington Books.

[31] Miceli, M.P., Near, J.P. and T.M. Dworkin, (2008), *Whistle-blowing in organizations*, New York: Routledge.

[32] Piccolo, S. and G. Immordino, (2017), "Organized Crime, Insider Information and Optimal Leniency," forthcoming in the *Economic Journal*.

[33] Rapp, G.C., (2012), "Mutiny by the Bounties? The Attempt to Reform Wall Street by the New Whistleblower Provisions of the Dodd-Frank Act," *Brigham Young University Law Review*, 1/2, 73-152. (Available at: http://digitalcommons.law.byu.edu/lawreview/vol2012/iss1/2).

[34] Rose, A.M., (2014), "Better Bounty Hunting: How the SEC's new whistleblower program changes the secretes fraud class action debate," *Northwestern University Law Review*, 108(4), 1235-1302.

[35] Schmolke, K.U. and V. Utikal, (2016), "Whistleblowing: Incentives and Situational Determinants," FAU - Discussion Papers in Economics, No. 09/2016, available at SSRN: https://ssrn.com/abstract=2820475 or http://dx.doi.org/10.2139/ssrn.2820475.

[36] Spagnolo, G., (2004), "Divide et Impera: Optimal Leniency Programmes," *CEPR Discussion Paper nr 4840.*

[37] Spagnolo, G., (2008), "Leniency and Whistleblowers in Antitrust," Ch.12 of P.Buccirossi (Ed.) *Handbook of Antitrust Economics*, Cambridge MA: M.I.T. Press.

[38] Stikeleather, B. (2016), "When do employers benefit from offering workers a financial reward for reporting internal misconduct?" *Accounting, Organizations and Society*, 52, 1-14.

[39] Transparency International, (2013), "Whistleblowing in Europe, Legal Protections for Whistleblowers in the EU"

[40] Wilde, J.H. (2017). "The Deterrent Effect of Employee Whistleblowing on Firms' Financial Misreporting and Tax Aggressiveness", *The Accounting Review*, In-Press.

[41] Zingales, L. (2004), "Want to Stop Corporate Fraud? Pay Off Those Whistle-Blowers," *The Washington Post*, January 18.