IMA Journal of Management Mathematics (2022) **33**, 289–314
https://doi.org/10.1093/imaman/dpab015
Advance Access publication on 10 May 2021

An evolutionary game theoretic model of whistleblowing behaviour in organizations

María J. Quinteros
Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Santiago, Chile 7941169 and Business School, Universidad Finis Terrae, Santiago, Chile 7501015

Marcelo J. Villena†
Universidad Técnica Federico Santa María, Santiago, Chile 7660251
†Corresponding author. Email: marcelo.villena@usm.cl

Mauricio G. Villena
Facultad Economía y Empresa, Universidad Diego Portales, Santiago, Chile 8581169

[Received on 15 March 2020; accepted on 30 March 2021]

Abstract

We present a theoretical model of corruption in organizations. Our specific focus is the role of incentives that aim to encourage whistleblowing behaviour. Corruption is modelled as a social norm of behaviour using evolutionary game theory. In particular, the dynamics of whistleblowing behaviour is captured using the replicator dynamics equation with constant and quadratic monitoring costs. We formally explore the local asymptotic stability of the equilibria. Our findings indicate that the traditional recommendations of the Beckerian approach are usually too expensive and/or unstable. We argue that an efficient mechanism for controlling corruption can be achieved by maintaining efficient salaries and imposing high rewards for whistleblowers when they detect wrongdoing. In the long term, employees can only be honest, or corrupt, or corrupt and whistleblowers; honest and whistleblowing behaviour will not coexist in the long run, since one of these two strategies is always dominated by the other.

Keywords: whistleblowing; corruption; social norms; evolutionary games; incentives.

1. Introduction

Corruption is a persistent feature in human societies, and similar to crime in general, it has always existed. Thus, corruption has always been an important concern for managerial teams in all kinds of organizations, private and public. Interestingly, most studies that have modelled corruption have focused on countries and governments, as well as on the use of punishment and surveillance to increase the expected cost of crime. See Rose-Ackerman (1975) for an early work in this topic and Aidt (2003) for a lucid survey of the economic literature.

Recently, whistleblowing has become an important monitoring mechanism in the wake of numerous corporate scandals involving accounting firms (Alleyne *et al.*, 2013). The business community and interested regulatory agencies are now calling for whistleblowing to be a prominent part of organizational culture.

For instance, in 2010, the US Securities and Exchange Commission (SEC) established the Dodd–Frank Whistleblower Program. The Dodd–Frank Act of 2010 included Section 21F (‘Securities Whistleblower Incentives and Protections’), which directs the SEC to issue monetary awards to individuals who voluntarily provide original information leading to successful enforcement of monetary sanctions of over $1 million. During 2017 alone, the SEC has ordered offenders to pay more than $975 million in total monetary sanctions in enforcement matters involving whistleblower information; this amount includes more than $671 million in disgorgement of profits acquired through securities fraud, which are in turn returned to investors who have been harmed. Whistleblowers are typically compensated in an amount equal to 10 to 30 percent of the monetary sanctions collected. In addition to establishing a reward program to encourage the submission of high-quality information, the SEC has implemented regulations prohibiting retaliation by firms against employees who have reported possible wrongdoing based on a reasonable belief that a possible securities violation has occurred, is in progress, or is about to occur.

Following the collapse of Arthur Andersen in the Enron scandal, whistleblowing has received considerable attention, primarily from the ethics literature; since then, business research on whistleblowing has also started to expand (Alleyne *et al.*, 2013). Nevertheless, from a management science perspective, research on whistleblowing behaviour has been rather scarce. In this context, this study aims to contribute by formally investigating the role of whistleblowing and its potential impact of whistleblowing on the persistence of corruption in organizations, focusing specifically on the role of incentives aimed at encouraging whistleblowing behaviour.

Whistleblowing can be formally defined as the disclosure by organization members (former or current) of illegal, immoral or illegitimate practices under the control of their employers to persons or organizations that may be able to effect action (Near & Miceli, 2016). A recent example of whistleblowing behaviour is that of a Monsanto executive who reported to the SEC improper accounting at the genetic engineering and pesticides producer Agrocorp. In an article on the subject, the New York Times reported that ‘the whistleblower got stymied within the company’ and he finally had to leave Monsanto during the course of the investigation. The whistleblower was reported to have stated that: *‘It’s really difficult when your company is doing something you know is wrong but you’ve got everybody around you saying it’s perfectly fine... the Monsanto culture is very tightknit. Everybody has stock options and everyone is financially at risk. So they go with the flow’*. After the SEC was informed that the company’s accounting practices for its Roundup herbicides were being used to overstate earnings, Monsanto agreed to settle for $80 million, and the Monsanto executive earned a $22 million award for blowing the whistle. While this case illustrates accusations of wrongdoing by an employee to his organization, there are few cases where an employee blows the whistle on wrongdoing by another employee.

From the business case discussed above, we can infer some distinctive elements about how whistleblowing functions in practice. For example, spontaneous whistleblowing behaviour can be difficult to initiate in practice because it implies costly activities and typically garners no economic rewards. The costs of whistleblowing include monitoring and transaction costs associated with actually reporting corrupt behaviour. These costs clearly depend on the number of agents within an organization that somehow favour the corrupt activities being committed. Indeed, transaction costs of reporting illicit activities can increase when more people within an organization support this type of behaviour. Thus, one way to approach the problem is to model corruption as a social norm of behaviour assuming that the illicit activities are somehow accepted by most agents as valid and common practices so that going against them means going against the majority.

In this study, we model corruption as a social norm of behaviour using elements of a classic game theory approach and elements of evolutionary game theory (EGT). Specifically, we use the concept of replicator dynamics—see for example Taylor & Jonker (1978), Vega-Redondo (1996), Hofbauer & Weibull (1996), Gintis (2000) and Villena & Villena (2004)—to formally explore the local asymptotic stability of the following types of behaviour within an organization: (i) honest behaviour, which implies that an employee does not receive any bribes from illicit activity; (ii) corrupt behaviour, in which a member of the organization does receive bribes from a corrupt relationship; and (iii) whistleblowing (honest enforcer) behaviour, in which an employee not only behaves honestly, but also monitors other people within the organization and reports them if they are behaving corruptly. Assuming that an organization includes three different population shares pursuing these three different behaviour types, we analyse which of these population shares will become stable within the organization in the long run. In other words, we formally explore the asymptotic stability of the non-corruption equilibrium, if all individuals behave honestly, the corruption equilibrium, if all individuals within the organization behave corruptly and mixed equilibria, where one share of the population is honest and/or whistleblower and another share is corrupt. To our knowledge, this is the first attempt to model whistleblowing behaviour using the concept of replicator dynamics to analyse the evolution of different types of behaviour within the organization, and evaluate whether apparently random states of disorder and irregularities are, in this particular case, governed by deterministic equations that are highly sensitive to initial conditions. In particular, we show some chaotic behaviour of the model, in other words dependence on initial conditions, and conditions under which our differential equation system converges to homeostasis if the parameters stay within certain limits.

Finally, our modelling strategy of whistleblowing behaviour allows us to consider additional control instruments for organizations to prevent corruption, going beyond the usual variables mentioned in the classic economic literature, namely, the wage rate, the monitoring system (probability of detection of a corrupt public official by an enforcing agency) and penalties for corrupt activities. In particular, our work adds to the analysis variables related to the environment for whistleblowing behaviour, namely, the role of the transaction costs associated with reporting a corrupt worker within an organization, the costs of monitoring an individual's activities and the economic incentives for whistleblowers that effectively detect and report a corrupt agent. No doubts, these economic variables can all be very important for organizations to consider as additional control mechanisms to promote whistleblowing behaviour and fight corruption. In this context, the potential effect of these instruments on the stability of corruption represents a very important issue to consider into the analysis.

Some questions we address in this research are: Will corrupt-whistleblowers contribute to break the stability of corruption or they will only add to the stability of this social norm of behaviour? How should economic instruments aiming at promoting whistleblowing behaviour be designed so they will effectively contribute to break the stability of corruption and not contribute instead to its stability within the organization? To the knowledge of the authors, no specific work squarely deals with the impact of economic incentives for whistleblowers in the stability of corruption. Consequently, we study in this article what would happen to the stability of corruption if these instruments are in place and if they are not.

The remainder of this paper is organized as follows. In Section 2, we present a literature review in which we emphasize previous theoretical works on whistleblowing behaviour and contextualize our contribution. Section 3 outlines whistleblowing policies in different countries. Section 4 describes our proposed model and its evolutionary dynamics. In section 5, we analyse the conditions under which the equilibria found are stable and under what conditions there is instability in the system. Section 6 presents numerical simulations in which we use elements of chaos theory to analyse our replicator dynamics system in order to evaluate how sensitive to initial conditions this system is. Finally, section 7 presents some concluding remarks and notes potential topics for future research.

2. Literature review

Gary Becker's seminal work entitled 'Crime and punishment: An economic approach' introduced the modelling of individual criminal behaviour as a response to incentives (Becker, 1968). An agent decides whether to commit a crime and how much crime to commit by comparing the expected costs and benefits of crime with those of alternative activities. In Becker's paper and its many successors (Block & Heineke, 1975; Ehrlich, 1973; Freeman, 1996), the main aim is to study to a certain extent how the probability of being caught, the magnitude of the punishment, the proceeds of criminal activity and the return to work (alternative to crime) would affect the level of crime. However, the Beckerian model is limited by the fact that agents face decisions at a moment in time where only auditing control methods exist, leaving out of the analysis the idea of peers monitoring peers within organizations.

Consequently, an extension to the literature on crime is the introduction of internal auditing of wrongdoing and anti-corruption social norms within organizations, for example, through mechanisms aimed at protecting and encouraging whistleblowing behaviour. In this context, some theoretical works have very recently started exploring whistleblowing behaviour from an economic perspective.

Soreide (2008) presents an economic framework aimed at explaining the potential reaction of multinationals to the lost of a contract because a competitor has offered a bribe. In particular, the author examines the impact of industry structure and institutional quality on the company's incentive to react against corruption. The paper assumes a homogenous output in a standard Cournot competition framework, considering in different contexts the potential costs and benefits of a firm' whistleblowing behaviour. The main conclusions of this study are, first, that firms will not react against a case of business corruption if that may disturb their opportunities to obtain cartel profits. Second, the more efficient the offender of the crime, the lower is the motivation for a potential whistleblower to react. Finally, a whistleblower reaction on corruption can cause other obstacles if there are connections between local politicians and firms in the given market.

There have been efforts before to model corruption using a game theory approach. Cerqueti & Coppier (2016) model the interaction between polluting firms, tax inspectors and politicians in a corrupted context. Interestingly, they found that, if the State fixes a high minimum incentive level, the only way to fight evasion is to implement a growing level of monitoring activity (hence, reducing the funds for the environment). Alternatively, a low minimum incentive threshold for the environmental inspector may lead to a more ethical route to fight evasion, which drives the policymaker towards an improvement of the fiscal performance. Cerqueti & Coppier (2018) model the relationship between bureaucratic corruption and political corruption considering the economic structure of a country. They found that the size of capital of a specific firm could influence the decision to bribe the bureaucrat.

Our work is in particular related to that of Heyes & Kapur (2008) in the sense that we also study the whistleblowing behaviour of individuals within an organization. They develop a behavioural model that adopts, in particular, the methods of behavioural law and economics. They put forward a general whistleblower 'motivation function', which relates whistleblowers' propensity to blow the whistle to the characteristics of the observed malfeasance and the enforcement environment. They gather evidence from sociology and psychology as to why employees blow the whistle on law-breaking employers, despite it not being in their narrowly defined self-interest. They identify three alternative 'schools of thought' and adapt the general motivation function to correspond to each school. They then characterize the optimal policy in each case. The main conclusions of this paper are, first, that optimal policy vary substantially between the cases. Second, the value of the information that whistleblowers bring to the enforcement agency, and what the agency will wish to do with that information, depends upon the motives ascribed to whistleblowers. Finally, in adjusting the enforcement instruments, attention has to be paid to the change induced in the flow of disclosures, in addition to the direct effect on compliance incentives. In this case also the quantitative and qualitative response will depend upon whistleblower motives.

Arce (2010) puts forward an evolutionary game-theoretic model to address three aspects of whistleblowing: ethical decision making, the duality of mutual accountability among cohorts in large organizations and role conflict between individual and organizational values. He finds an equilibrium condition relating the treatment of whistleblowers to the punishment of violators, which can be used to characterize the relationship between the treatment of whistleblowers and those of violators. Moreover, a direct link is forged between the pay-offs and rewards for actions within an inspection game and the frequencies of these strategies for the population in question. In this case violations are minimized when whistleblowers are treated as heroes. This is consistent with the idea that whistleblowers should be rewarded for serving the public good.

Brianzoni *et al.* (2019) analysed the behaviour of non-compliant firms using a dynamic setting where honest or dishonest behaviour prevails in society at any given time. They conclude that for a society to address dishonest behaviour it can either set up a high level of monitoring, or ‘build’ a high level of ‘inner honesty’ in society.

Acemoglu & Jackson (2017) study the relationship between social norms and the enforcement of laws, being their key argument that many laws are ineffective, in part because they conflict with prevailing social norms, making private agents unwilling to cooperate with law enforcement, for example, by whistleblowing, while at the same time effective laws successfully change social norms, significantly increasing their potency. In their model, individuals choose a behaviour, e.g. corruption, and then are matched uniformly with another agent. Utility depends negatively on the average behaviour of other agents and on the mismatch between the behaviours of the two partners. A law is a cap on behaviour and a law-breaker, when detected, pays a fine and has her behaviour forced down to the cap. Incentives to break the law depend on social norms because detection has to rely, at least in part, on private cooperation and whistleblowing. Law-abiding agents have an incentive to whistleblow because this will reduce their partner’s behaviour, ameliorating the mismatch. When laws are in conflict with norms so that many others are breaking the law, anticipating little whistleblowing, each agent has further incentives to also break the law.

While the self-reinforcing, dynamic, nature of corruption, implying that the greater the number of people who adhere to corrupt activities is, the more persistent corruption becomes, has been extensively reviewed in the literature, see for instance Lui (1986), Cadot (1987), Sah & Stiglitz (1988), Andvig & Moene (1990), Murphy *et al.* (1991), Acemoglu & Jackson (2017), Baker *et al.* (1994) and Tirole (1996), to our knowledge there are no previous works squarely analysing this characteristic of corruption in the context of whistleblowing behaviour.

Our approach to modelling social norms differs from these authors since it is based on EGT. The basic characteristic of this approach is that it does not assume optimizing behaviour per se, but does retain the idea that individuals adjust their behaviour in response to persistent differentials in material incentives. Thus, although economic agents do pursue individual material pay-offs, which, in these models, represent evolutionary success, they are not always in a position to obtain straightaway the pay-offs an optimizing agent would obtain. This is because social norms of behaviour restrict the course of action of individuals in such a way as to prevent them from adjusting their behaviour towards the optimal strategy immediately (it takes time to change a social norm followed by the majority of the population).

However, if this situation persists in time, some individuals will start adopting the more efficient strategy and will, therefore, receive a higher pay-off than the rest of the population. In the long run, the rest of the population will start imitating this more profitable course of action. Thus, the incumbent social norm will be replaced by this new, more successful strategy that, in time, will be adopted as the new norm of behaviour in the population. In this sense, evolutionary models can be interpreted as models of bounded-rationality, where individuals learn about the game on a trial-and-error basis and where more efficient behaviour, in evolutionary terms, tends to be imitated. This alternative approach based on EGT allows us to analyse the impact of different initial population shares ascribing to whistleblowing behaviour and the relevance the parameters when explaining the stability (or instability) of corruption.

3. Whistleblower policies in practice

Several countries have laws designed to encourage whistleblowers to come forward with her complaints/concerns as part of their anti-corruption and freedom of information laws (Banisar, 2011).

In this context, the recent approved European Whistleblower Protection Directive draws attention for its strength and completeness, pointing to the relevance or this public policy against corruption. Indeed, on December 2019, the European Parliament approved a 'Whistleblower Protection Directive' containing broad free speech protections for whistleblowers in both the public and the private sectors for all member states of the European Union(see Directive (EU) 2019/1937 of the European Parliament and of the Council on the Protection of Persons Who Report Breaches of Union Law–Whistleblower Directive). The Directive prohibits direct (or indirect) retaliation against public and private sector employees. The Directive's protections apply to employees and those who assist them, including civil society organizations and journalists who report on their evidence. It establishes equal rights for whistleblowers from the national security sector who challenge the denial or withdrawal of their security clearances. In addition, whistleblowers are protected against criminal prosecution and company lawsuits for damages resulting from their reporting, and are provided with psychological support to cope with the stress caused by harassment.

Without doubts, the USA is one the prime exponents of the whistleblower policy. It has been claimed that whistleblowing tradition started in the USA in the 1770s with Benjamin Franklin leaking letters on the Hutchinson affair. In the USA, legal protections vary depending upon the case matter and the state where the case arises. For instance, the Whistleblower Protection Act of 1989 is a federal law that protects federal whistleblowers who work for the government and report the possible existence of unlawful behaviour such as mismanagement, gross waste of funds, abuse of authority, potential danger to public health and safety. The Directorate of Whistleblower Protection Program (DWPP) of the United States Department of Labor's Occupational Safety and Health Administration (OSHA) is in charge of the investigation of retaliation against whistleblowers under several federal statutes.

In the UK, the Public Interest Disclosure Act (PIDA) of 1998 provides protection to workers making disclosures in the public interest and allows such individuals to claim compensation for victimization following such disclosures. Further protection was afforded by The Enterprise and Regulatory Reform Act 2013 (ERRA), which came into force in July 2013. In Australia, there exists the Corporations Act 2001 in order to encourage whistleblowers to come forward and protect them, which gives people legal rights and protections as whistleblowers. From 1 July 2019, the whistleblower protections in the Corporations Act have been expanded to provide greater protections for whistleblowers.

In Canada, the Office of Public Servants Disclosure Protection Act (PSIC) of 2007 provides safe mechanism enabling public servants and the general public to disclose wrongdoings committed in the public sector. The Office protects public servants from retaliation who have disclosed wrongdoing.

4. The model

We consider an organization in which employees display the following types of behaviour: honest, corrupt and whistleblower.

An employee receives a wage of w for her work. Therefore, an honest employee gets paid w. If an employee commits illicit acts, she receives an amount β for her illicit acts. However, if the corrupt employee is caught, she is fired, and she can get an alternative job at a wage of $w_0 \geq 0$ and will pay to the actual firm a penalty of $f \geq 0$. When $w > w_0$ we assume that the employee's wage is higher than the market wage. In this context, we are talking of an efficiency wage strategy from the organization point of view. The firm may have some whistleblower employees who monitor others and denounce them when caught in illicit acts. Whistleblowers can detect wrongdoing behaviour with a probability of $\Theta > 0$. However, as discussed in the introduction, it is difficult for whistleblowing behaviour to arise spontaneously since it is a costly activity. The costs associated with whistleblowing include monitoring costs of $m > 0$ and transaction costs of $\tau > 0$ associated with actually reporting wrongdoing.

We assume that whistleblowing employees are honest. Given that, it is clear that no pay-off maximizer agent will become a whistleblower since this activity only lowers her pay-off regardless of the behaviour of the rest of the employees. In fact, in this case, the whistleblowing strategy is strictly dominated by the honest strategy since $w > w - \tau\Theta - m$, so no employee has the incentive to become a whistleblower. This, in turn, implies that no employee is deterred from corruption by the threat of whistleblowing behaviour from her colleagues.

Because whistleblowing behaviour is a dominated strategy, we introduce a reward denoted by $\sigma \geq 0$ for monitoring the work done by their colleagues and reporting illicit acts. In addition to whistleblowing behaviour among employees, a mechanism–for example, auditing–can also detect wrongdoing behaviour with probability $\theta \geq 0$.

Considering a total population of n employees, a proportion of p_1 is honest, p_2 is corrupt and p_3 is made up of whistleblowers ($p_1 + p_2 + p_3 = 1$). Each of the $p_1 n$ honest employees receives the pay-off associated with honest behaviour, a wage of $\pi_1 = w$. The proportion of corrupt employees $p_2 n$ receive the expected pay-off associated with wrongdoing activities, $\pi_2 = (1 - \theta - p_3\Theta)(w + \beta) + (\theta + p_3\Theta)(w_0 - f)$, assuming that corrupt acts can be caught by auditing with probability θ or by any of the whistleblowers with probability $p_3\Theta$. If she is not caught, she receives her wage and the bribery payment, and if she is caught, she is fired and must work in another place and pay the penalty. Finally, each of the $p_3 n$ whistleblowers receives the expected pay-off of $\pi_3 = w + p_2\Theta(\sigma - \tau) - m$, which is associated with her wage plus the reward for blowing the whistle less the costs associated with monitoring and reporting. The average pay-off of all employees is $\bar{\pi} = p_1\pi_1 + p_2\pi_2 + p_3\pi_3$.

4.1 *The dynamics of whistleblowing behaviour*

We model whistleblowing behaviour within an organization using the replicator dynamic equation, formulating a simple model of evolution and biased learning in games in which successful strategies spread by natural selection among a population. These model has been used in biology and also in economics to study evolution. of populations In our context, we propose that employers are not completely committed to just one way of behaving. Rather, several possible ways of behaving are present (honest, corrupt, whistleblower) in their minds simultaneously. Which type of behaviour predominate, and which are given less attention, depends on the experiences of the individual. The change in population behaviour can be analogous to biological evolution.

To formalize the replicator dynamics equation, consider an evolutionary game with q pure strategies and stage game pay-off π_{ij} for any i-player who meets any j-player. If $p = (p_1, ..., p_q)$ is the frequency of each type of population, the expected pay-off for the i-player is then $\pi_i(p) = \sum_{j=1}^{q} p_j \pi_{ij}$, and the average pay-off in the game is $\bar{\pi}(p) = \sum_{j=1}^{q} p_j \pi_i$. The replicator dynamics equation for our model is then given by

$$\frac{dp_i}{dt} = \dot{p_i} = p_i[\pi_i(p) - \bar{\pi}(p)] \qquad \forall i = 1(honest),\ 2(corrupt) \tag{4.1}$$

The replicator equation 4.1 expresses the idea that strategies grow among population if they do better than average. The strategies that do best grow fastest.

We model the evolution of employees behaviour assuming (1) constant monitoring costs i.e. regardless of the proportion of corrupts, the costs are the same, and (2) following Becker (1968), who suggests that $\frac{\partial m}{\partial p_2} > 0$ and $\frac{\partial^2 m}{\partial p_2^2} > 0$, we propose quadratic monitoring costs, assuming then, $m(p_2) = ap_2^2 + bp_2 + m$.

The replicator dynamics equation in the context of our model is represented by the following two differential equations:

The replicator dynamics model when monitoring costs are constant is given by

$$\begin{aligned} \dot{p_1} = p_1 &\left[m(1-p_1) + p_2\Theta(1-p_1-p_2)(w+\beta-\sigma+\tau)\right] + \\ &p_1p_2\left[(p_1+p_2)(w_0-f) - \beta(1-\theta) + \theta(w-w_0+f) - m + f - w_0\right] \end{aligned} \tag{4.2}$$

$$\begin{aligned} \dot{p_2} = p_2 &\left[p_1(f-m-w_0) - \Theta(1-p_1)(w+\beta) + \beta(1-\theta) - \theta(w-w_0+f)\right] + \\ &p_2[m+w_0-f] + p_2^2\left[\theta(w-w_0+f) - \beta(1-\theta) + 2(f-w_0) - m\right] + \\ &p_2^2\left[(p_1+p_2)(\Theta(\sigma-w-\tau-\beta) + w_0 - f) + \Theta(2w+2\beta-\sigma+\tau)\right] \end{aligned} \tag{4.3}$$

The replicator dynamics model when monitoring costs are quadratic becomes

$$\begin{aligned} \dot{p_1} = -[&ap_2^3 + [(\beta-\sigma+w+\tau)\Theta + p_1a - a + b + f - w_0]p_2^2 + [(p_1-1)(\beta-\sigma+w+\tau)\Theta + \\ &(b+f-w_0)p_1 + (1-\theta)\beta - (f+w-w_0)\theta - b + m - f + w_0]p_2 + m(p_1-1)]p_1 \end{aligned} \tag{4.4}$$

$$\begin{aligned} \dot{p_2} = -(&ap_2^3 + ((\beta-\sigma+w+\tau)\Theta + p_1a - a + b + f - w_0)p_2^2 + \\ &(((\beta-\sigma+w+\tau)p_1 - 2\beta + \sigma - 2w - \tau)\Theta + \\ &(b+f-w_0)p_1 + (1-\theta)\beta - (f+w-w_0)\theta - b + m - 2f + 2w_0)p_2 - \\ &(p_1-1)(w+\beta)\Theta + (m-f+w_0)p_1 + (\theta-1)\beta + (f+w-w_0)\theta - m + f - w_0)p_2 \end{aligned} \tag{4.5}$$

By solving the system $\dot{p_1} = 0, \dot{p_2} = 0$ for both, constant and quadratic monitoring costs, we found five solutions in which the system converges (fixed points of the system). The first equilibrium is when all employees in the long term become whistleblowers, represented by $p_1^* = p_2^* = 0$ and $p_3^* = 1$. A second equilibrium is when all employees in the long term become honest ($p_1^* = 1$ and $p_2^* = p_3^* = 0$). The third equilibrium is when all employees in the long term are corrupt ($p_1^* = p_3^* = 0$ and $p_2^* = 1$). The fourth case is when employees in the long term are either corrupt or whistleblowers, represented by $p_1^* = 0$, $p_2^* > 0$ and $p_3^* > 0$. The last equilibrium is when the three types of employees coexist: honest employees, corrupt employees and whistleblowers ($p_1^* > 0$, $p_2^* > 0$ and $p_3^* > 0$).

5. Analysis of the equilibria

In section 4, we calculated the fixed points of the replicator dynamics equation using constant and quadratic monitoring costs.

In this section, we are interested in examining the stability of the system since stability is what tells us where the model converges under certain given conditions to find the variables that encourage employees to behave honestly. Our model is non-linear, and we use the (indirect) Lyapunov's linearization method to study the stability of each fixed point. Let p_1^* and p_2^* be the fixed points of the original system, then $\dot{p}_i = p_1 \frac{\partial \dot{p}_i(p_1^*, p_2^*)}{\partial p_1} + p_2 \frac{\partial \dot{p}_i(p_1^*, p_2^*)}{\partial p_2}$, with $i = 1, 2$. To study stability, we only need to find the eigenvalues of the system presented and if $Re(\lambda_i) < 0$, then (p_1^*, p_2^*) is asymptotically stable.

The Jacobian matrix of the model with constant monitoring costs is given by

$$J_1 = \begin{bmatrix} \begin{array}{c} p_2\Theta(2p_1 + p_2^2 - 1)(\sigma - w - \tau - \beta) \\ +p_2(2p_1 + p_2^2)(w_0 - f) \\ -p_2\beta(1-\theta) + p_2\theta(w - w_0 + f) \\ +p_2(f - m - w_0) + m(1 - 2p_1) \end{array} & \begin{array}{c} p_1\Theta(1 - p_1 - 2p_2)(w + \beta - \sigma - \tau) \\ +p_1(p_1 + 2p_2)(w_0 - f) - p_1\beta(1-\theta) \\ +p_1\theta(w - w0 + f) + p_1(f - m - w_0) \end{array} \\ \begin{array}{c} p_2^2(\Theta(w + \tau + \beta - \sigma) + f - w_0) \\ +p_2(\Theta(w + \beta) + f - m - w_0) \end{array} & \begin{array}{c} p_2\Theta(2p_1 + 3p_2)(\sigma - w - \tau - \beta) \\ +p_2(2p_1 + 3p_2)(w_0 - f) + 2p_2\Theta(2w + 2\beta - \sigma + \tau) \\ -2p_2\beta(1-\theta) + \theta f(1 + 2p_2) - (1 - p_1)(f - m - w_0) \\ -\Theta(1 - p_1)(w + \beta) + \beta(1-\theta) - \theta(1 - 2p_2)(w - w_0) \end{array} \end{bmatrix} \tag{5.1}$$

The Jacobian matrix of the model with quadratic monitoring costs is given by

$$J_2 = \begin{bmatrix} \begin{array}{c} p_2^2(\Theta(\sigma - \tau - w - \beta) - 2p_1 a + a - b - f + w_0) \\ +p_2(-\Theta(2(w + \beta - \sigma + \tau))(p1 - \frac{1}{2}) + \\ 2p_1(w_0 - b - f) + \beta(\theta - 1) + \\ \theta(f + w - w_0) + b + f - m - w_0) \\ -2mp_1 + m - ap_2^3 \end{array} & \begin{array}{c} -2p_1(\frac{\Theta}{2}(p_1 + 2p_2 - 1)(w + \beta - \sigma + \tau) \\ +\frac{3ap_2^2}{2} + p_2(ap_1 - a + b + f - w_0) + \\ \frac{p_1}{2}(f + b - w_0) + \frac{\beta}{2}(1-\theta) \\ -\frac{\theta}{2}(f + w - w_0) + \frac{m - b - f + w_0}{2}) \end{array} \\ \begin{array}{c} p_2(-ap_2^2 + p_2(\Theta(\sigma - \tau - w - \beta) - b - f + w_0) \\ +\Theta(\beta + w) + f - m - w_0) \end{array} & \begin{array}{c} 3p2^2(\Theta(\sigma - \tau - w - \beta) \\ -ap_1 + a - b - f + w_0) + \\ p_2(\Theta(2p_1(\sigma - \tau - w - \beta) \\ +4\beta - 2\sigma + 4w + 2\tau) + 2p_1(w_0 - b - f) \\ +2\beta(\theta - 1) + 2\theta(f + w - w_0) + \\ 2b + 4f - 2m - 4w_0) + \Theta(p_1 - 1)(w + \beta) + \\ p_1(f - m - w_0) + \beta(1-\theta) + \theta(-f - w + w0) \\ -f + m + w_0 - 4ap_2^3 \end{array} \end{bmatrix} \tag{5.2}$$

In what follows, we will study only the case of constant monitoring costs, since both models converge to the same steady states.

5.1 *On the stability of the whistleblower*

By solving the characteristic polynomial of the Jacobian (equation 5.1) with fixed points $p_1^* = p_2^* = 0$, the eigenvalues are $\lambda_1 = m$ and $\lambda_2 = -\Theta\beta - \Theta f - \Theta w + \Theta w_0 - \beta\theta - f\theta - \theta w + \theta w_0 + \beta + m$. It is clear that the solution $p_1^* = p_2^* = 0$ is unstable since $\lambda_1 = m > 0$, meaning that the solution is unstable regardless of the values of the other parameters.

PROPOSITION 1. It is a dominant strategy to behave honestly over whistleblowing when there are no corrupt employees.

Proof. The pay-off for an honest employee is $\pi_h = w$. The pay-off for a whistleblower employee when there is no corrupt population is $\pi_w = w - m$. Since $w > w - m$, it is clear that when there is no wrongdoing, the population will all behave honestly. □

If all the population is whistleblower, and hence honest, the whistleblower employees will perceive the cost of monitoring without any expected reward, since corruption doesn't exist in this setting. Obviously, this is not an evolutionary stable strategy.

5.2 *On the stability of the honest*

PROPOSITION 2. The honest equilibrium ($p_1^* = 1, p_2^* = 0$) is locally asymptotically stable if and only if $w > w_0 - \beta - f + \frac{\beta}{\theta}$.

Proof. solving the characteristic polynomial of the Jacobian when $p_1^* = 1, p_2^* = 0$, gives us $\lambda_1 = -m$ and $\lambda_2 = -\beta\theta - f\theta - \theta w + \theta w_0 + \beta$. Since $\lambda_1 < 0$, we only need to impose the condition $\lambda_2 < 0$ in order to get a stable fixed point, which is true when $w > w_0 - \beta - f + \frac{\beta}{\theta}$. □

Honest population, in order to be an evolutionary stable strategy, needs to comply that $w > w_0 - \beta - f + \frac{\beta}{\theta}$. This means that high penalty rates f and a high probability of being detected by auditing θ imply significant expected costs of dismissal, so the honest equilibrium will be stable and hard to break. By contrast, the higher the amount of the bribery β is, the more easily our honest equilibrium can be broken. On the other hand, the design of a successful auditing control and monitoring mechanism can be both difficult and costly. In practice, it may imply a very low θ. If $\theta \to 0$ implies that the wage w necessary to avoid corruption becomes too high and actually very hard to pay, and then, if there is no monitoring, salaries alone cannot avoid corruption and a firm with only honest employees is never achieved, reaffirming the importance of a potential equilibrium including a population of whistleblowers.

5.3 *On the stability of the corrupt*

An environment of corruption can be developed whenever $p_1^* = 0, p_2^* = 1$. By solving the characteristic polynomial of the Jacobian the eigenvalues are: $\lambda_1 = \theta(\beta + f + w - w_0) - \beta$ and $\lambda_2 = \Theta(\sigma - \tau) + \theta(\beta + f + w - w_0) - \beta - m$. The conditions for employees becoming corrupt in the long run depend on Θ.

PROPOSITION 3. The corruption equilibrium is asymptotically stable if and only if $\Theta < \frac{m}{\sigma - \tau}$ and $w < w_0 - \beta - f + \frac{\beta}{\theta}$

The proof of proposition 3 is as follows. To obtain stability, we must impose $\lambda_1 < 0$ and $\lambda_2 < 0$. For λ_1, the condition is true if $w < w_0 - \beta - f + \frac{\beta}{\theta}$. Since $\Theta < \frac{m}{\sigma - \tau}$, when $w < w_0 - \beta - f + \frac{\beta}{\theta}$ implies that $\lambda_2 < 0$. ■

The corrupt equilibrium will be an evolutionary stable strategy, and hence hard to break when: (i) the probability a whistleblower detects a corrupt employee is low, that is, $\Theta < \frac{m}{\sigma - \tau}$, and/or the probability of being detected by auditing is rather small, and/or the penalty for being caught in a corrupt activity, f, is low and (ii) increasing either the amount of bribery, β, or the wage if dismissed, w_0.

In fact, if $\theta \to 0$, the condition for the local asymptotic stability of the corruption equilibrium is always satisfied, which implies that whenever corruption becomes a common practice, it will be very difficult to stop.

By contrast, an increase in salaries may break the stability of corruption. This argument dates back to Becker & Stigler (1974) who noted that high salaries can be used to control corruption since they increase the cost of dismissal and, therefore, make employees more reluctant to accept bribes. Nevertheless, paying high salaries can be very expensive for firms and does not ensure that corruption will be reduced in all situations. In the next sections, we will analyse the dynamics of this Becker's recommendation, highlighting some instability issues that would prevent the use of high salaries to control corruption.

5.4 *On the stability of corrupt and whistleblowers*

This means that in the long run, $p_1^* = 0$ and $p_2^* = B$, $B \in (0, 1]$. In this environment, a firm has employees that are either corrupt or whistleblowers.

PROPOSITION 4. The equilibrium of corrupt and whistleblower employees is asymptotically stable if and only if $\Theta > \frac{m}{\sigma - \tau}$, $w < w_0 - \beta - f + \frac{\beta}{\theta}$, $i - kw < 0$ and $j - lw < 0$

Proof of proposition 4: First notice that if $\Theta < \frac{m}{\sigma - \tau}$ the replicator dynamic is always asymptotically stable and converges to either all honest or all corrupt employees. If $w < w_0 - \beta - f + \frac{\beta}{\theta}$, the model converges to the corruption equilibrium when $\Theta < \frac{m}{\sigma - \tau}$ (from proposition 3). If $w > w_0 - \beta - f + \frac{\beta}{\theta}$, the model always converges to the honest equilibrium, in particular when $\Theta < \frac{m}{\sigma - \tau}$, so stability of corrupts and whistleblowers occurs when $\Theta > \frac{m}{\sigma - \tau}$ and $w < w_0 - \beta - f + \frac{\beta}{\theta}$ and whenever the eigenvalues are negative. Before solving the characteristic polynomial of the Jacobian, let's define the following notation:

$i = -\Theta(f + \tau - w_0 + \beta - \sigma)B^2 + (\Theta(f + \tau - w_0 + \beta - \sigma) + (f - w_0 + \beta)\theta - m - \beta)B + m$

$j = -3\Theta(f + \tau - w_0 + \beta - \sigma)B^2 + ((4f + 2\tau - 4w_0 + 4\beta - 2\sigma)\Theta + (2f - 2w_0 + 2\beta)\theta - 2m - 2\beta)B + (-f + w_0 - \beta)\Theta + (-f + w_0 - \beta)\theta + m + \beta$

$k = B(\Theta B - \Theta - \theta)$ $l = 3\Theta B^2 + (-4\Theta - 2\theta)B + \Theta + \theta$

The solution of the characteristic polynomial of the Jacobian (equation 5.1) is then $\lambda_1 = i - kw$ and $\lambda_2 = j - lw$. This means that this equilibrium is stable when $i - kw < 0$ and $j - lw < 0$. ■

Although it is never good for an organization to reach the case of only corrupt employees, the case of only honest employees can be extremely expensive and very difficult to achieve, in particular when the bribe β is high or when the probability of auditing detecting corruption, θ, is very low. In these cases, for the organization, it may be more efficient to control corruption but not eliminate it all together. A firm with only honest employees may have salaries $w > w_0 - \beta - f + \frac{\beta}{\theta}$. When the firm has whistleblower and corrupt employees, in comparison to the salaries paid when the firm only has honest employees, salaries are reduced. However, the firm has the additional expected costs of corruption $p_2\beta - f(\theta + p_3\Theta) + p_3\Theta\sigma$.

Corollary 5.1. When $p_1^* = 0$ and $p_2^* = B = 1$, the equilibrium of only corrupt employees is asymptotically stable if and only if $\Theta > \frac{m}{\sigma-\tau}$ and $w < w_0 - \beta - f + \frac{\beta}{\theta} + \frac{m}{\theta} - \Theta\frac{\sigma-\tau}{\theta}$.

The proof of corollary 5.1 is as follows. To obtain stability, we must impose $\lambda_1 < 0$ and $\lambda_2 < 0$. Given $B = 1 \Rightarrow \lambda_1 = \beta\theta + f\theta + \theta w - w_0\theta - \beta$ and $\lambda_2 = \Theta\sigma - \Theta\tau + \beta\theta + f\theta + w\theta - w_0\theta - \beta - m$. For λ_1, the condition is true if $w < w_0 - \beta - f + \frac{\beta}{\theta}$; however, for λ_2, since $\Theta > \frac{m}{\sigma-\tau}$, $\lambda_1 < \lambda_2$, and then we must impose $\Theta\sigma - \Theta\tau + \beta\theta + f\theta + w\theta - w_0\theta - \beta - m < 0$, which is equivalent to $w < w_0 - \beta - f + \frac{\beta}{\theta} + \frac{m}{\theta} - \Theta\frac{\sigma-\tau}{\theta}$. ■

Assuming $\sigma > \tau$, when the probability of catching a corrupt employee is low, $\Theta < \frac{m}{\sigma-\tau}$, then wages must be high to avoid corruption. The wage condition is relaxed to some extent (by $\frac{m}{\theta} - \Theta\frac{\sigma-\tau}{\theta}$) when the probability $\Theta > \frac{m}{\sigma-\tau}$. In other words, cost savings of the whistleblowing policy is proportional to $\frac{m}{\theta} - \Theta\frac{\sigma-\tau}{\theta}$.

5.5 *On the stability of honest, corrupt and whistleblowers*

An environment where honest, corrupt and whistleblowing employees coexist occurs when $p_1^* = A$, $p_2^* = B$, with $A, B > 0$ and $A + B < 1$. This equilibrium is not stable. The model is asymptotically stable under the circumstances described above for an environment of all honest employees, an environment of all corrupt employees or an environment in which employees are either corrupt or whistleblowers. Intuitively, when the reward for blowing the whistle is higher than the cost, then being honest is dominated by whistleblowing. By contrast, if the costs of whistleblowing are greater than the reward, then whistleblowing is dominated by a strategy of being honest. Thus, in the long term, employees can be only honest, only corrupt, or corrupt and whistleblowers; honest and whistleblowing behaviour won't coexist in the long run, since one of these two strategies is always dominated by the other.

6. Numerical simulations

In this section, we show numerical examples of the model to illustrate some of our theoretical results. In particular, we are interested in simulating the parameters associated with wrongdoing β, in order to study under which conditions it is worth having a whistleblowing culture, the reward σ, which is the new control parameter against wrongdoing, the wage w to study which level of wage is necessary to avoid a culture of wrongdoing, the fine f, in order to study if the Beckerian model is fulfilled, and finally the probabilities θ and Θ to understand the level of monitoring effort necessary to combat corruption.

Note that in all the experiments, except the first one, we use the wage as numeraire. In other words, we set the value of the wage to 1 and for each experiment we made the value of a parameter vary. For instance, when we vary the bribe β between 0.1 and 4, it implies that it varies in terms relative to the wage between 10 and 400%.

6.1 *Simulations for the wage w*

We simulate values for $w \in [1, 8.8]$. Figure 1 shows the proportion of each type of employees an organization might have. Panel (a) shows what would happen in an environment where the wrongdoing gains β go from 50 to 6.25% of the wage, while in panel (b) β goes from 200 to 25% of the wage, all other parameters fixed.

In Fig. 1(a), proposition 2 is satisfied whenever $w > 2$, as shown graphically.

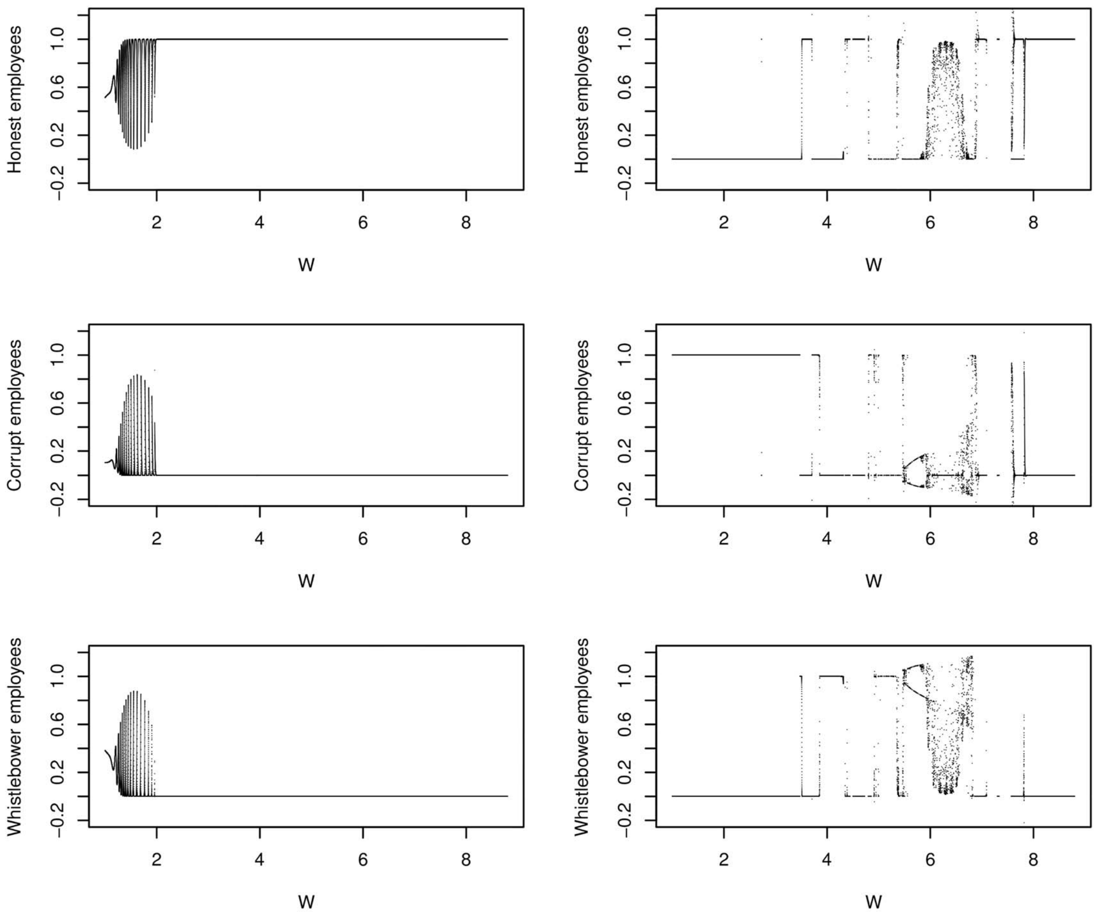

FIG. 1. The bifurcation cascade for honest, corrupt and whistleblower employees as a function of the wage. Parameters: $w_0 = 0.1$, $f = 0.1$, $\theta = 0.2$, $\Theta = 0.35$, $\tau = 0.2$, $m = 0.1$. (a) Left: $\beta = 0.5$, (b) Right: $\beta = 2$.

Proposition 3 or corollary 5.1 are never satisfied, and then an organization will not reach an equilibrium of only corrupt employees. When $w < 2$ there is instability in the system. In Fig. 1(b), proposition 2 is satisfied when $w > 8$. Proposition 3 is never satisfied since $\Theta > \frac{m}{\sigma-\tau}$. However, corollary 5.1 is satisfied when $w < 3.6$, which means that whenever the wage is less than 180% of the bribe, corruption will be the norm. The organization will present unstable behaviour (the eigenvalues will be one positive and one negative) when salaries are between 3.6 and 8, suggesting that high salaries alone are not enough to control corruption.

In Fig. 2 we show the proportion of each type of employee for $w = 1$, $w = 4$ and $w = 8$ in the long term. In both panels we see that when $w = 8$ the organization has only honest employees. When the bribe is set to 2 (see panel (b)) and $w = 5$, although it is a high wage, it fails to control corruption, with the proportion of each type of employee being unstable in the long term. Finally, when $w = 1$ and the bribe is low (situation in panel (a)), there is instability, however with a higher proportion of honest and

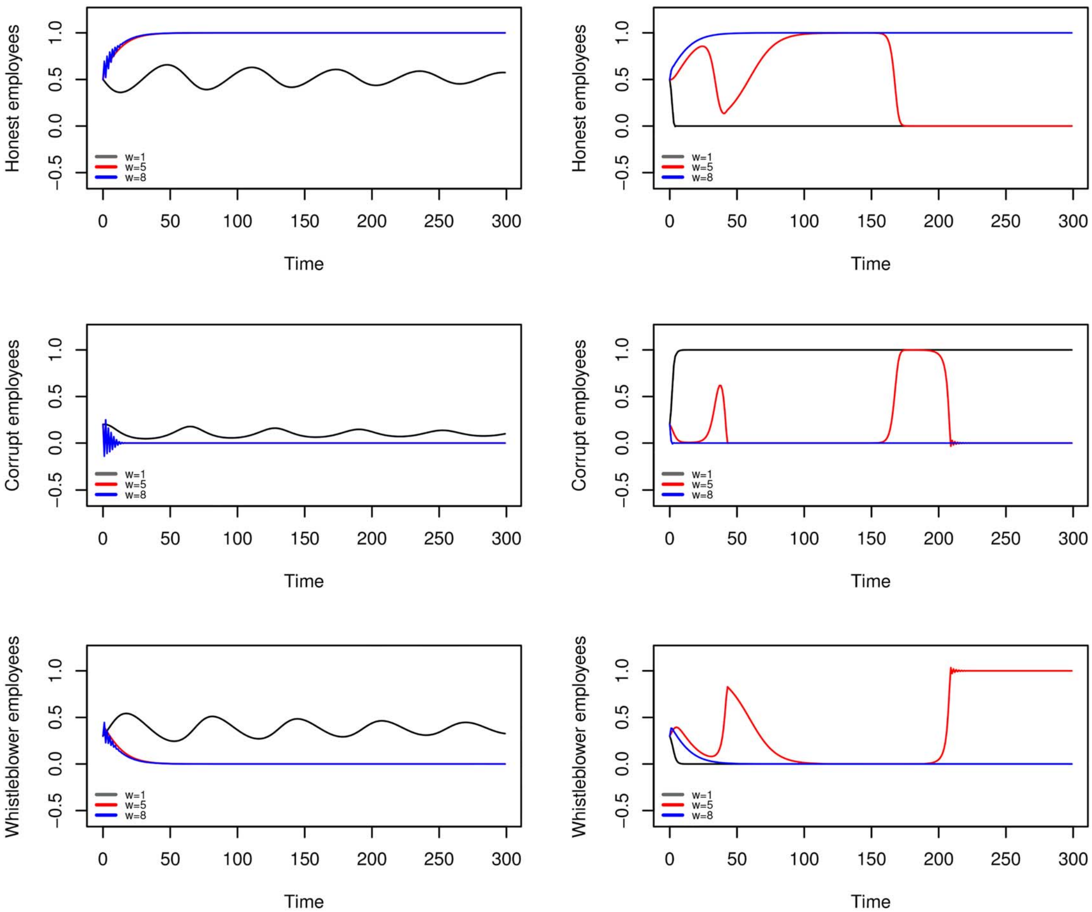

FIG. 2. Time series of system for $w = 1$, $w = 5$ and $w = 8$. Parameters: $w_0 = 0.1, f = 0.1, \theta = 0.2, \Theta = 0.35, \tau = 0.2, m = 0.1$. (a) Left: $\beta = 0.5$, (b) Right: $\beta = 2$.

whistleblower over corrupt employees. In panel (b), since bribe is rather high, the organization will end up having only corrupt employees.

6.2 *Simulations for the fine f*

We simulate values for $f \in [0.01, 4]$. Figures 3 and 4 show what would happen with the organization setting different values for the fine. Panel (a) shows the dynamic when $\beta = 0.8$ and panel (b) when $\beta = 2$ in both figures.

Figure 3 shows graphically the results of our propositions. In particular, for panel (a), proposition 2 is satisfied whenever $f > 2.3$. Since $\Theta > \frac{m}{\sigma-\tau}$, proposition 3 is never satisfied. In addition, corollary 5.1 is also never satisfied, since the condition is that $f < -2.1$, and fine is always positive. This means, under this setting, the organization doesn't converge to having only corrupt employees. Eigenvalues, when $f < 2.3$ takes values greater than 1, and then, the proportion of each type of employee become unstable. For panel (b), proposition 2 is satisfied whenever $f > 7.1$. As in the previous case, proposition 3 is

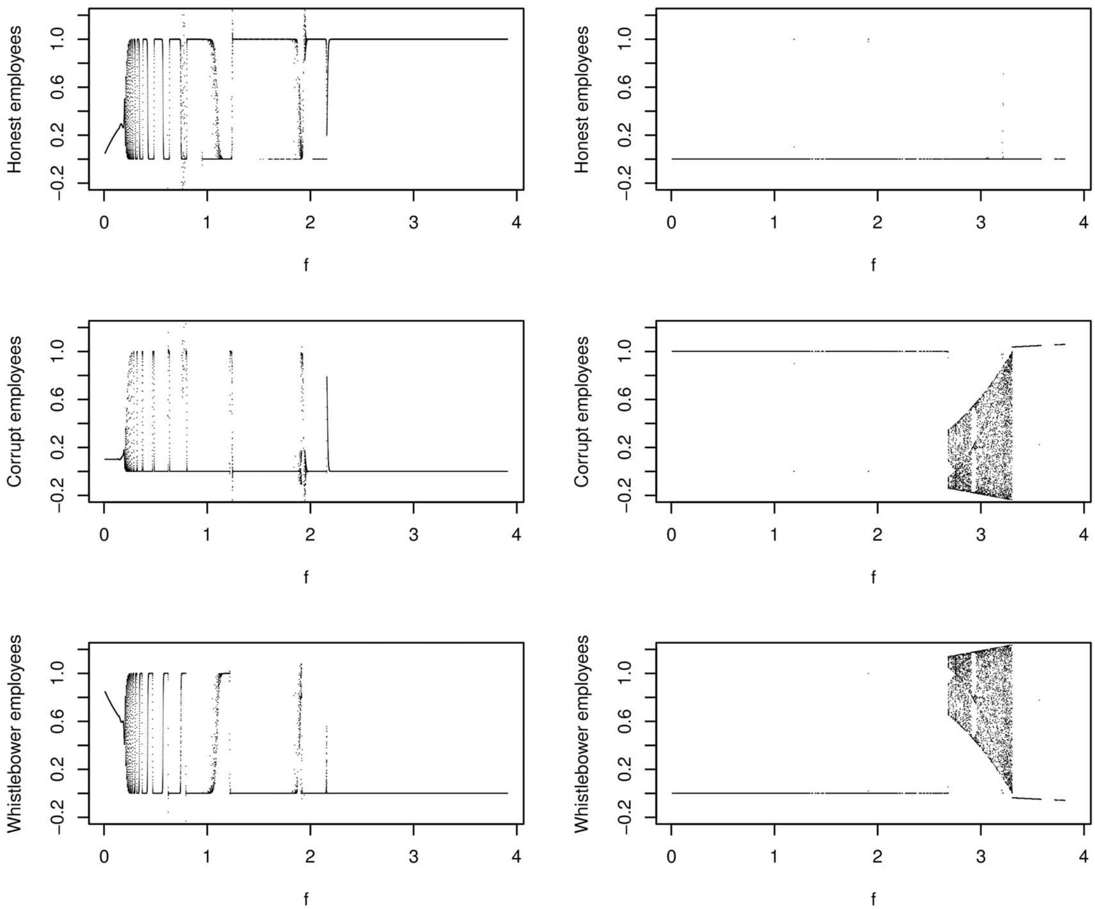

FIG. 3. The bifurcation cascade for honest, corrupt and whistleblower employees as a function of the penalty. Parameters: $w = 1$, $w_0 = 0.1$, $\theta = 0.2$, $\Theta = 0.35$, $\tau = 0.2$, $\sigma = 3$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

never fulfilled. However, corollary 5.1 is true whenever $f < 2.7$ and the eigenvalues associated with the characteristic polynomial are negative, which occurs for almost all values of $f < 2.7$. When the fine is relatively high ($f \in (2.7, 7.1)$), the organization will present an unstable proportion of each type of employees, in which case, high fines do not fulfil the purpose of deterring corruption, as pointed out by Polinsky & Shavell (1979).

In Fig. 4, we show some simulations in the long term for $f = 0.8$, $f = 3$ and $f = 4$, in which we exemplify what we discussed above, showing instability in panel (a) for $f = 0.8$ and in panel (b) for $f = 2$ and $f = 4$.

6.3 *Simulations for the earnings per wrongdoing β*

We simulate values for $\beta \in [0.1, 4]$. Figure 5 shows the existence of instability when earnings per wrongdoing are similar to the reward for reporting corruption. Panel (a) shows what would happen in an environment where wage is set to 1 and the reward for blowing the whistle is high, 300% of the wage,

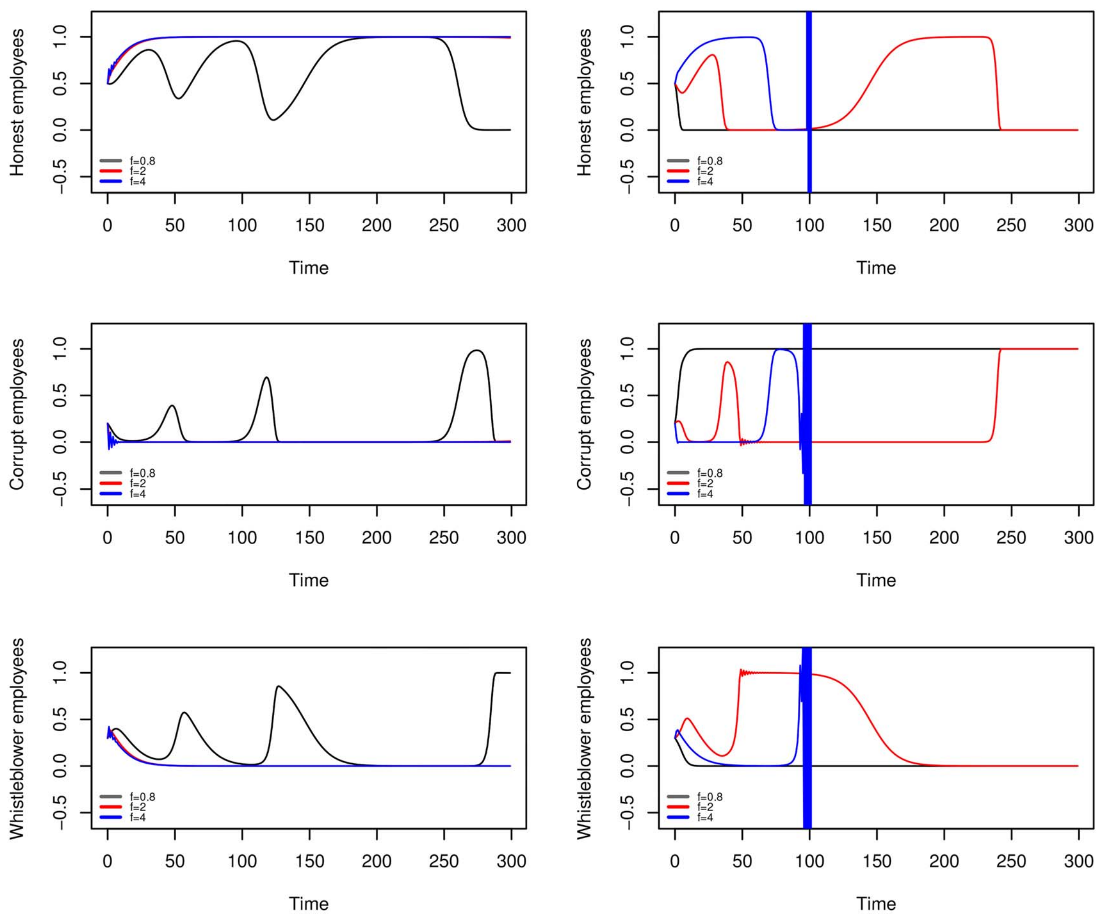

FIG. 4. Time series of system for $f = 0.8, f = 3$ and $f = 4$. Parameters: $w = 1$, $w_0 = 0.1$, $\theta = 0.2$, $\Theta = 0.35$, $\tau = 0.2$, $\sigma = 3$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

while panel (b) represents an organization where the wage is very high ($w = 10$) and the reward σ is low compare with that level of wage (5% of w).

For Fig. 5(a), proposition 2 is satisfied whenever $\beta < 0.25$. Since $\Theta > \frac{m}{\sigma-\tau}$, proposition 4 and its corollary 5.1 imply whistleblower and corrupt employees whenever $\beta \in (0.25, 0.825)$, as shown in panel (a). For $\beta > 0.825$ there is a large range for β values where only corrupt employees exist. However, when β becomes large enough, the eigenvalues associated with the characteristic polynomial start taking some values higher than 0, presenting unstable behaviour. For Fig. 5(b), proposition 2 is satisfied whenever $\beta < 2.5$. Since $\Theta < \frac{m}{\sigma-\tau}$, proposition 3 is satisfied whenever $\beta > 2.5$, presenting only unstable behaviour around $\beta = 2.5$, as shows panel (b).

The reward for whistleblowing controls corruption when the value of this reward is high in relation to the offenses. Figure 5 shows that when the reward for whistleblowing is low (panel (b)), it fails to control corruption even though the wage is high. In panel (a), for earnings per wrongdoing of the order of one-third of the reward blowing the whistle, the organization is able to control corruption.

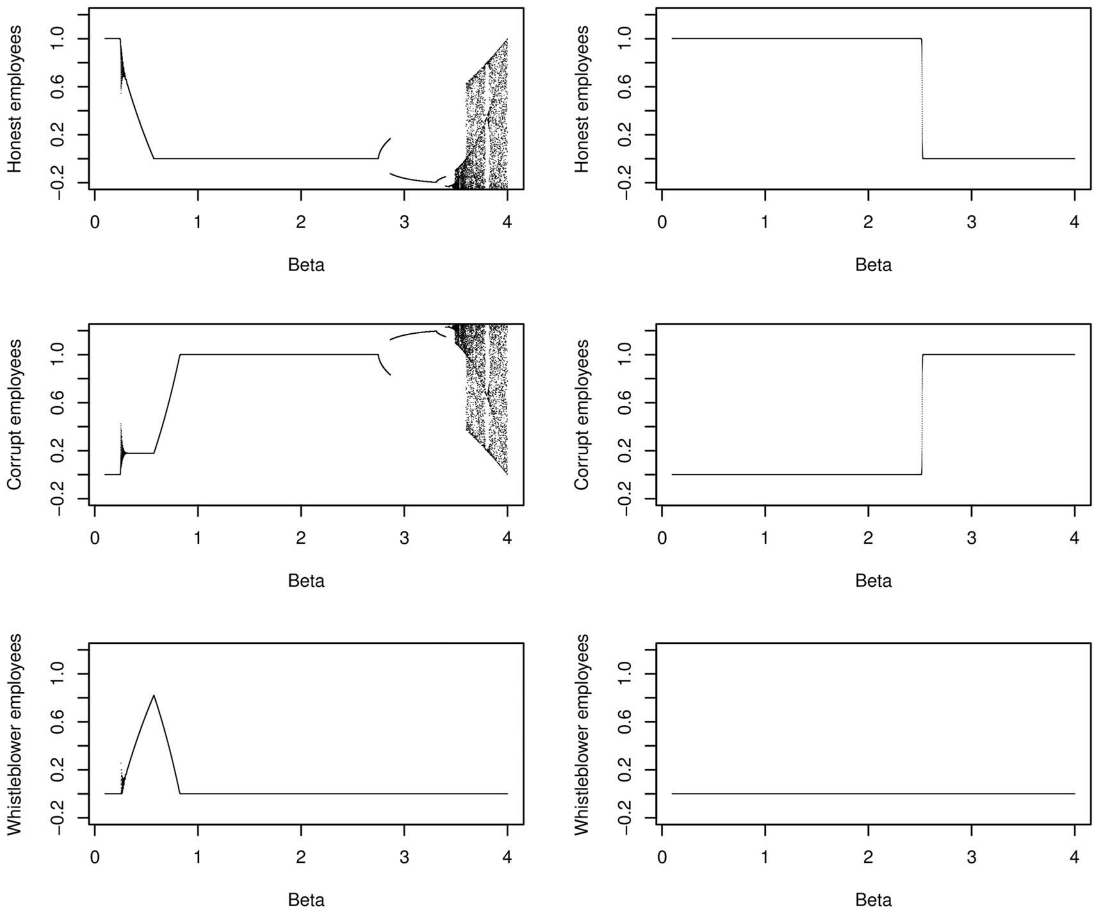

FIG. 5. The bifurcation cascade for honest, corrupt and whistleblower employees as a function of the earnings per wrongdoing. Parameters: $w_0 = 0.1, f = 0.1, \theta = 0.2, \Theta = 0.2, \tau = 0.2, m = 0.1$. (a) Left: $w = 1, \sigma = 3$, (b) Right: $w = 10, \sigma = 0.5$.

In Fig. 6, we simulate the proportion of each type of employee in the long term for $\beta = 0.8$, $\beta = 3$ and $\beta = 4$. Panel (a) shows that for high amounts of bribe ($\beta = 3$ and $\beta = 4$), the system becomes uncontrolled, and then it is hard to predict the proportion of each type of employee in the organization. For $\beta = 0.8$ the system is stable. In panel (b), for $\beta = 3$ and $\beta = 4$, the organization converges to only corrupt employees, not fulfilling high salaries with the purpose of minimizing corruption.

6.4 *Simulations for the auditing probability of detecting wrongdoing θ*

We simulate values for $\theta \in [0.01, 0.75]$. Figures 7 and 8 show the development of an organization by varying the auditing probability of detecting wrongdoing in two scenarios, $\beta = 0.8$ and $\beta = 2$.

Figure 7(a) shows that proposition 2 is satisfied whenever $\theta > 0.\overline{4}$, as simulations in Fig. 8(a) for $\theta = 0.56$ and $\theta = 0.6$ show. The equilibrium of only corrupt employees is never satisfied. Firstly, since $\Theta > \frac{m}{\sigma - \tau}$ proposition 3 is never met. Secondly, to satisfy corollary 5.1 $\theta < 0$, which does not satisfy the property that θ is a probability. Proposition 4 is satisfied whenever $B \in (0.0982, 0.5541)$ and $\theta \in$

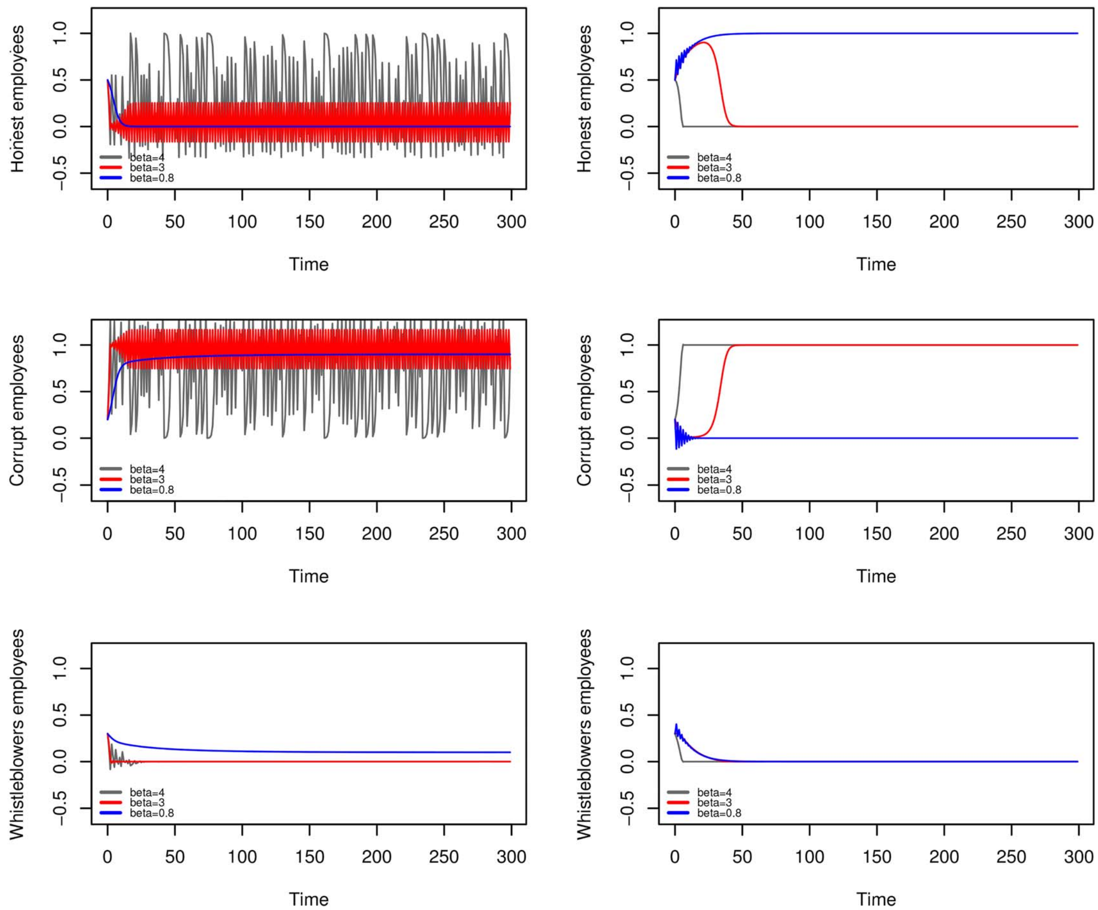

FIG. 6. Time series of system for $\beta = 4$, $\beta = 3$ and $\beta = 0.8$. Parameters: $w_0 = 0.1$, $f = 0.1$, $\theta = 0.2$, $\Theta = 0.2$, $\tau = 0.2$, $m = 0.1$. (a) Left: $w = 1$, $\sigma = 3$, (b) Right: $w = 10$, $\sigma = 0.5$.

$\left(\frac{0.00\overline{5}(105B^2-124B+27)}{1-2B}, -\frac{0.02\overline{7}(7B^2-25B+2.)}{B}\right)$ or $B > 0.5541$ and $\theta < \frac{0.00\overline{5}(105B^2-124B+27)}{1-2B}$. Figure 7(b) shows that proposition 2 is satisfied whenever $\theta > 0.\overline{6}$. Proposition 3 is satisfied whenever $\theta < 0.37\overline{3}$. There is an unstable behaviour among the employees whenever $\theta \in (0.37\overline{3}, 0.\overline{6})$. When earnings per worngdoing are high, as shown in Fig. 7, where β is twice the wage, high levels of monitoring fails to control corruption, contradicting the Beckerian results, which point to the fact that the higher the level of monitoring, the lower the levels of corruption achieved.

Figure 8(b) shows unstable proportion of each type of employee in the long term for $\theta = 0.56$ and $\theta = 0.6$, and only corrupt employees in the long term when $\theta = 0.2$, consistent with the propositions.

6.5 *Simulations for the probability a whistleblower detects wrongdoing* Θ

We simulate values for $\Theta \in [0.01, 0.8]$. Figures 9 and 10 show that for our simulations, we obtain that relatively low values of Θ do not deter corruption. Moderate probability of detecting wrongdoing

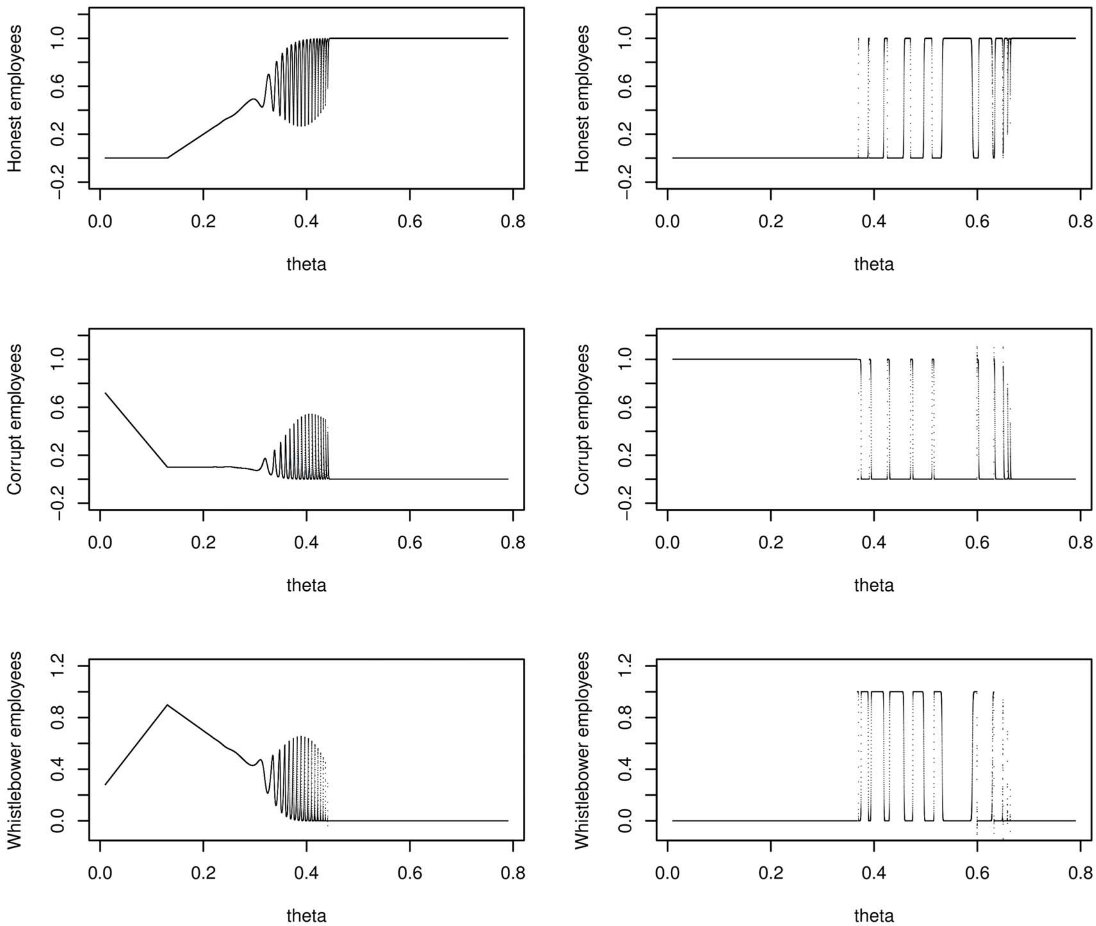

FIG. 7. The bifurcation cascade for honest, corrupt and whistleblower employees as a function of the auditing monitoring probability of detecting wrongdoing. Parameters: $w = 1$, $w_0 = 0.1$, $f = 0.1$, $\Theta = 0.35$, $\tau = 0.2$, $\sigma = 3$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

conduct to an steady state with a share of whistleblower and corrupts. For rather high values of Θ, the model becomes unstable, specially when earnings per wrongdoing are rather low, making it impossible to predict what type of employees the organization would have in the long term.

For all the simulations in Fig. 9, proposition 2 is never satisfied, and then an organization with only honest employees will not happen. The condition in panel (a) for having only honest employees is $w > 3.2$ and in panel (b), we obtain that $w > 8$, which are extremely high salaries comparing with the other parameters of the model. When $\Theta < 0.1929$, proposition 3 is satisfied, and only corrupt employees will be the norm. Proposition 4 is satisfied in the following cases:

- $B \in [0, 0.1307]$ and $\Theta > \frac{0.02(27B-5)}{B(B-1)}$
- $B \in (0.1307, 0.5409]$ and $\Theta < \frac{2.7(2B-1)}{15B^2+8B-9}$

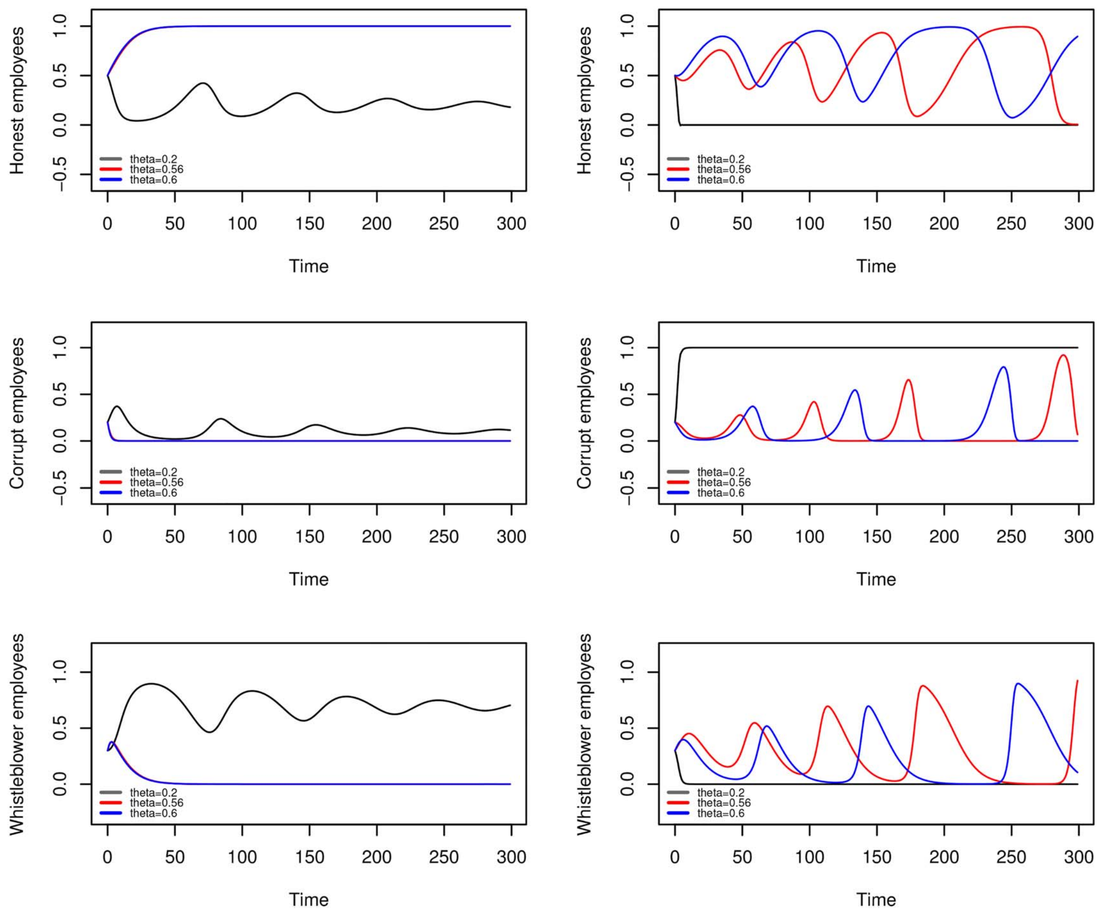

FIG. 8. Time series of system for $\theta = 0.2$, $\theta = 0.56$ and $\theta = 0.6$. Parameters: $w = 1$, $w_0 = 0.1$, $f = 0.1$, $\Theta = 0.35$, $\tau = 0.2$, $\sigma = 3$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

- $B \in (0.5409, 0.5525]$ and $\Theta > \frac{0.02(27B-5)}{B(B-1)}$
- $B \in (0.5525, 1]$ and $\Theta \in \left(\frac{0.02(27B-5)}{B(B-1)}, \frac{2.7(2B-1)}{15B^2+8B-9}\right)$

When earnings per wrongdoing are relatively low, it is sufficient to control corruption with a probability Θ greater, but close to $\frac{m}{\sigma-\tau}$. If it is too easy to detect wrongdoing (Θ close to 1), then the system becomes unstable. One explanation for this situation is as follows: since it is easy to find wrongdoing and the reward for them is low, then we will be in a situation where crime does not pay and the whistleblowers will find corruption easily, and therefore there will be periods of only honest employees. However, since there are no whistleblowers, it will become attractive again to commit offenses and therefore will increase the proportion of corrupt employees and at the same time, since there is more corruption, it will become attractive to be a whistleblower again. In contrast, if the earnings per wrongdoing are relatively high, the probability Θ must be high to bring corruption under control.

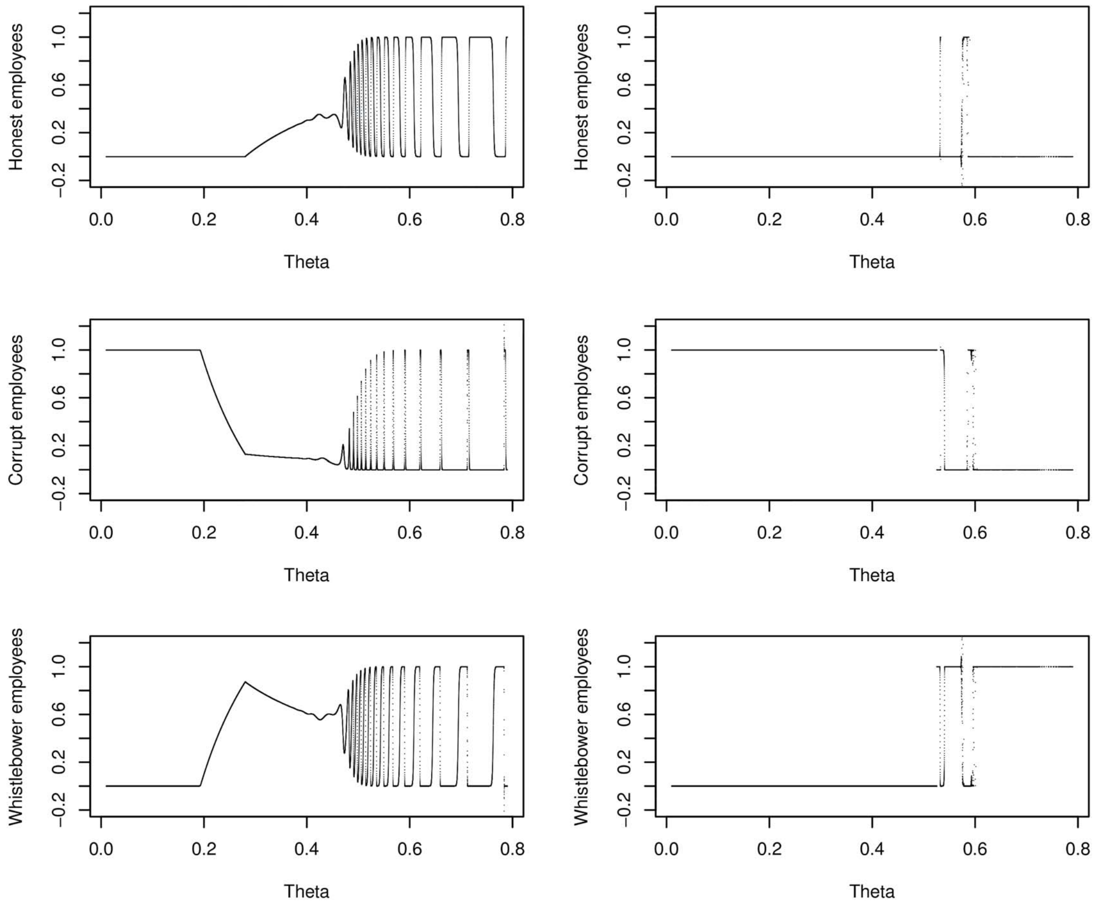

FIG. 9. The bifurcation cascade for honest, corrupt and whistleblower employees as a function of the probability a whistleblower detects wrongdoing. Parameters: $w = 1$, $w_0 = 0.1$, $f = 0.1$, $\theta = 0.2$, $\tau = 0.2$, $\sigma = 3$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

6.6 *Simulations for the reward for blowing the whistle σ*

For σ—the reward for blowing the whistle—we simulate values for $\sigma \in [0.1, 8]$. Figure 11 shows the results of those simulations. Panel (a) shows what would happen in an environment where the earning associated with wrongdoing β is moderated (80% of wage), while panel (b) represents an organization where the earning associated with wrongdoing is high in comparison with the wage (200% of w).

For Fig. 11, proposition 2 is never satisfied, since for both cases—results in panel (a) and (b)—$w < w_0 - \beta - f + \frac{\beta}{\theta}$, which implies that there would never be only honest employees in this organization.

In panel (a), $\Theta < \frac{m}{\sigma - \tau}$ whenever $\sigma \in (0.2, 0.486)$, where proposition 3 is satisfied, and then, in the long term will be only corrupts employees. Additionally, corollary 5.1 also implies that there only will be corrupt employees in the organization whenever $\sigma < 1.743$. In other words, when the incentives to blow the whistle are rather low (less than 1.743), incurring the costs of monitoring and reporting becomes unattractive. This panel shows instability when the bribe β is low. In panel (b), corollary 5.1 is satisfied

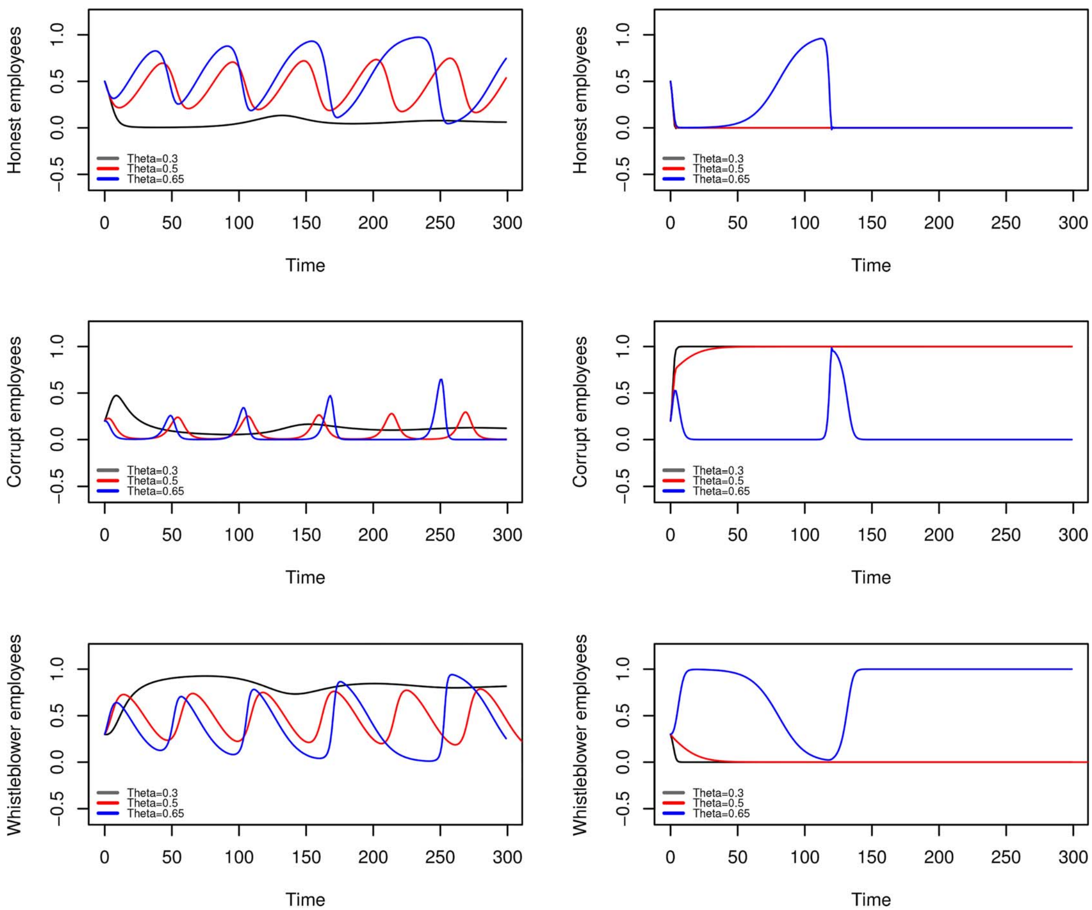

FIG. 10. Time series of system for $\Theta = 0.3$, $\Theta = 0.5$ and $\Theta = 0.65$. Parameters: $w = 1$, $w_0 = 0.1$, $f = 0.1$, $\theta = 0.2$, $\tau = 0.2$, $\sigma = 3$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

whenever $\sigma < 4.486$, implying the same as in the previous case, low values of σ are unattractive to report wrongdoing. However, even when β is high, when the reward σ is large enough, the organization could be described as 'healthy', having low levels of corruption, without the need for increasing the salaries of all employees.

In Fig. 12, we simulate the proportion of each type of employee for $\sigma = 2.5$, $\sigma = 4$ and $\sigma = 6$ in the long term. Panel (a), as pointed out above, shows that the system becomes uncontrolled for all three scenarios. However, for $\sigma = 6$ and $\sigma = 4$, oscillations have low wavelength range, with a considerable larger proportion of honest and whistleblowers over corrupt employees. In panel (b), for $\sigma = 2.5$ and $\sigma = 4$, the organization converges to only corrupt employees. However, when $\sigma = 6$, the organization will have more than 72% of whistleblower employees and around 27% of corrupt employees.

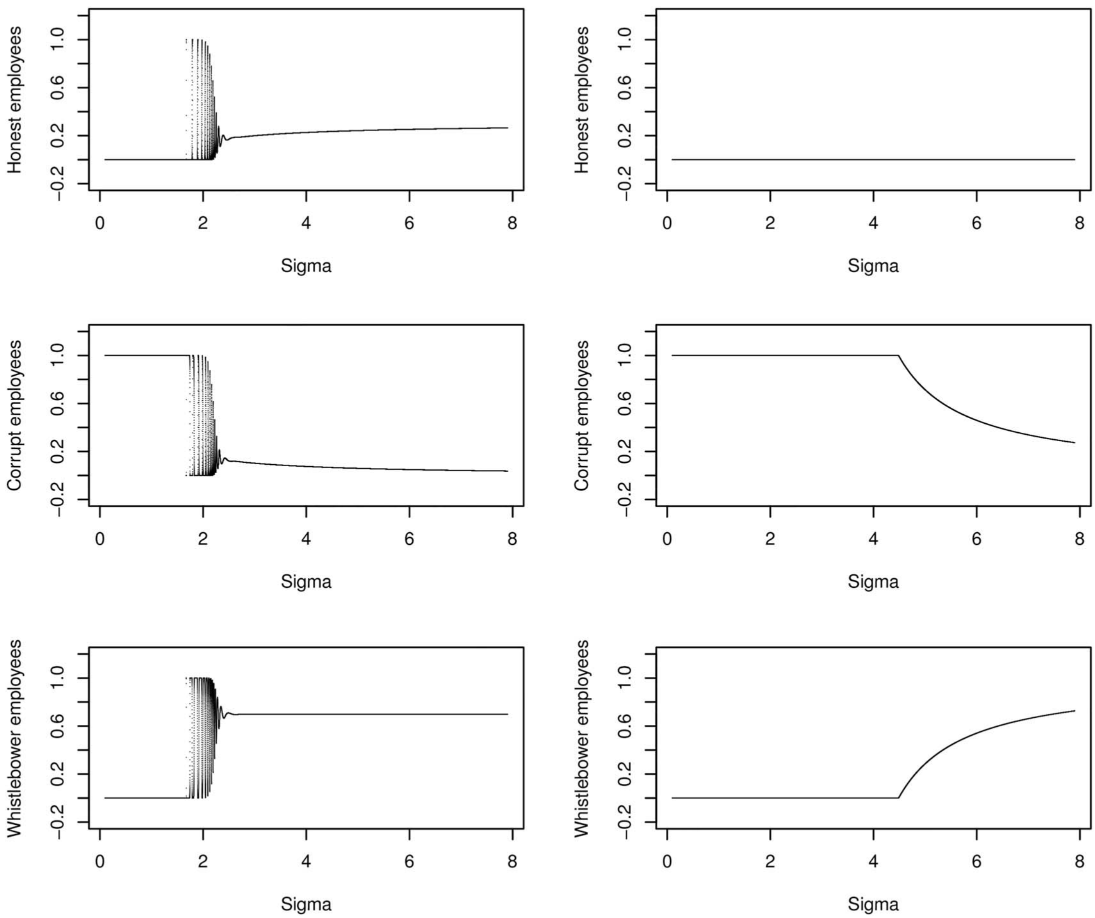

FIG. 11. The bifurcation cascade for honest, corrupt and whistleblower employees as a function of the reward for whistleblowing. Parameters: $w = 1$, $w_0 = 0.1$, $f = 0.1$, $\theta = 0.2$, $\Theta = 0.35$, $\tau = 0.2$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

7. Concluding remarks

In this paper, we have developed and solved a theoretical model to examine corruption in organizations. Our modelling efforts focused specifically on the role of incentives aimed at encouraging whistleblowing behaviour. In particular, we model corruption as a social norm of behaviour using EGT. In the model, we include the effect of the classical variables mentioned in the economic literature to control corruption in an organization, such as wage, auditing monitoring system and fines. Our work also adds to the analysis variables related to the environment for whistleblowing behaviour, namely, the role of the transaction costs associated with reporting a corrupt worker within an organization, the costs of monitoring individual's activities and the economic incentives for whistleblowers that effectively detect and report a corrupt agent.

We model the dynamics of whistleblowing behaviour using the replicator dynamic equation with constant and quadratic monitoring costs. Both models converge to the same five long-term steady states,

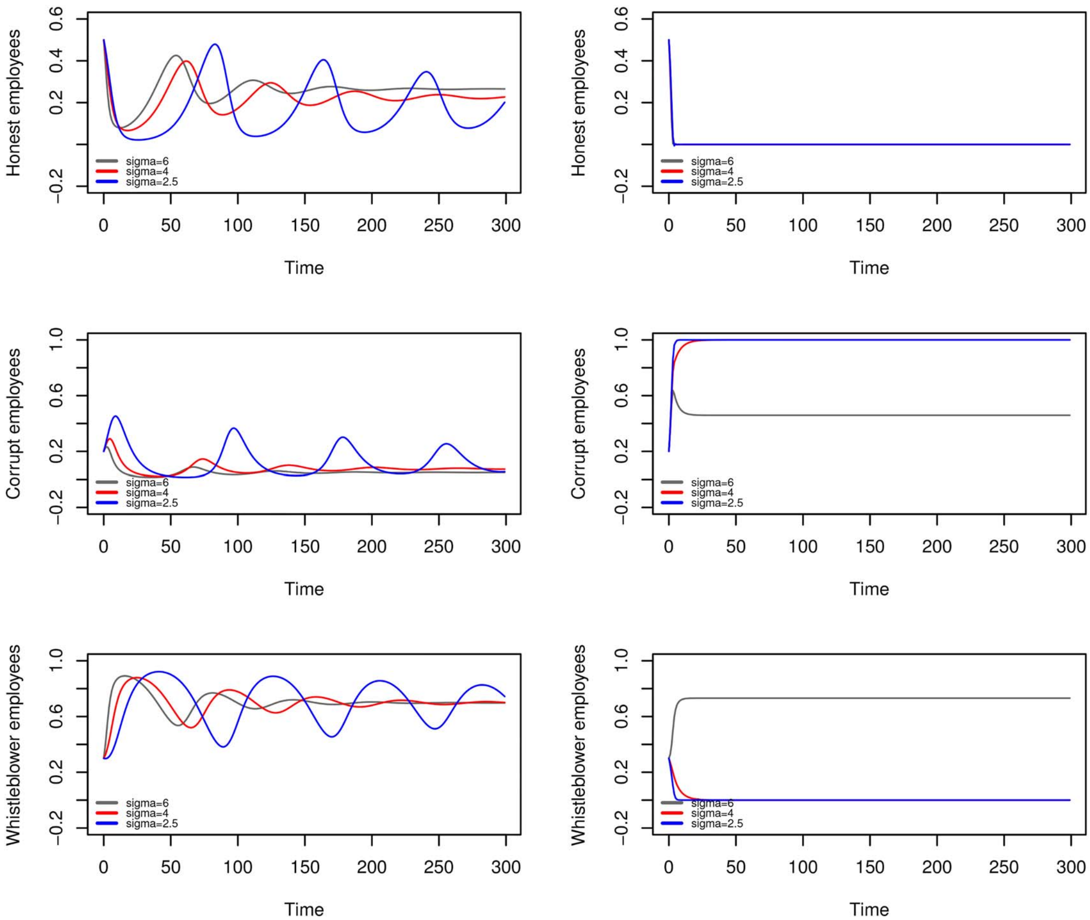

FIG. 12. Time series of system for $\sigma = 6$, $\sigma = 4$ and $\sigma = 2.5$. Parameters: $w = 1$, $w_0 = 0.1$, $f = 0.1$, $\theta = 0.2$, $\Theta = 0.35$, $\tau = 0.2$, $m = 0.1$. (a) Left: $\beta = 0.8$, (b) Right: $\beta = 2$.

pointing to the robustness of the model: (i) all employees become honest, (ii) all corrupt, (iii) one share of employees becomes corrupt and the other share whistleblowers, (iv) all whistleblowers and (v) a share of each type of employee. From these five equilibria, the first three are stable. First, the equilibrium of only honest employees is stable when salaries are set high. Second, the equilibrium of only corrupt employees is an evolutionary stable strategy when salaries are set very low. Third, the equilibrium of corrupt and whistleblowers is the optimal strategy when salaries are set lower than salaries for only honest employees, but the probability with which a whistleblower detects wrongdoing is relatively high. Fourth, the all-whistleblowers equilibrium is clearly unstable behaviour since the whistleblower will incur in monitoring costs and since no corruption will be detected, it never will receive the reward for blowing the whistle. Finally, a share of each type of employees is unstable behaviour, because when the reward for blowing the whistle is higher than the cost, then being honest is dominated by being whistleblower. By contrast, if the costs of whistleblowing behaviour are greater than the reward, then whistleblower behaviour is dominated by the being honest strategy. Thus, in the long term, employees can behave as only honest, only corrupt, or corrupt and whistleblowers; honest and whistleblowing behaviour won't coexist in the long run, since one of these two strategies is always dominated by the other.

The main conclusions of the paper are the following:

- As pointed out in some previous literature, just by having high salaries we may be able to control corruption and even reach a state without corruption. However, reaching such goal may become very expensive and in practice impossible to implement given its chaotic behaviour in a wide range of possible wage levels, and then, this mechanism alone is not capable of controlling corruption, it even generates instability.
- High penalties, as a mechanism to control corruption, are only effective when the earnings associated with wrongdoing are relatively low compared to wage. The higher the value of the earnings per wrongdoing, the less likely high penalties will deter corruption in the long term, and then not fulfilling its purpose as a mechanism to control corruption.
- Whistleblowing, as a mechanism to control wrongdoing, is only relevant under the existence of auditing monitoring. If the probability of detection of wrongdoing with an auditing mechanism is close to cero then, in the long term, all employees eventually will behave corruptly. For a firm to have only honest employees may be very hard and expensive to achieve, mainly when the probability to detect corruption by an auditing agency is low.
- The probability of detecting corruption through whistleblower employees is only relevant in a limited range of all possible values that it may take. If this probability is very low, there will be no incentives to become a whistleblower; moreover, if the probability of detecting corruption through auditing is also low, the organization will end up with only corrupt employees. By contrast, if the probability of detecting corruption through whistleblowers is very high, that is, it is easy to find corrupt employees, the organization will have instability in the shares of each type of employee.
- The award for reporting corruption, as a mechanism to control corruption, must be high to encourage whistleblowing behaviour. If this award is high enough, it is not necessary to incur in high salaries for all employees to combat corruption.

In summary, our main findings point to the fact that the traditional recommendations of the Beckerian approach are usually either too expensive and/or unstable (chaotic), while an efficient mechanism for controlling corruption within an organization can be achieved by maintaining efficient salaries and imposing high rewards for whistleblowers when they detect wrongdoing. In this manner, in the long term, corrupt behaviour will be minimized and the reward for blowing the whistle will become a cheap way to deter crime. An interesting further work in this line of research could be to include the possibility that whistleblowers are also corrupt, even though the complexity of such a potentially realistic assumption could be intractable.

References

Acemoglu, D. & Jackson, M. (2017) Social norms and the enforcement of laws. *Journal of the European Economic Association*, **15**, 245–295.

Aidt, T. (2003) Economic analysis of corruption: a survey. *The Economic Journal*, **113**, F632–F652.

Alleyne, P., Hudaib, M. & Pike, R. (2013) Towards a conceptual model of whistle-blowing intentions among external auditors. *The British Accounting Review*, **45**, 10–23.

ANDVIG, J. & MOENE, K. (1990) How corruption may corrupt. *Journal of Economic Behavior & Organization*, **13**, 63–76.

ARCE, D. G. (2010) Corporate virtue: Treatment of whistle blowers and the punishment of violators. *European Journal of Political Economy*, **26**, 363–371.

BAKER, G., GIBBONS, R. & MURPHY, K. J. (1994) Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*, **109**, 1125–1156.

BANISAR, D. (2011) Whistleblowing: International standards and developments. *Corruption and transparency: Debating the frontiers between state, market and society* (I. SANDOVAL ed). Washington, DC: World Bank-Institute for Social Research, UNAM.

BECKER, G. S. (1968) Crime and punishment: An economic approach. *The economic dimensions of crime*. London: Springer, pp. 13–68.

BECKER, G. S. & STIGLER, J. (1974) Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, **3**, 1–18.

BLOCK, M. & HEINEKE, J. (1975) A labor theoretic analysis of the criminal choice. *The American Economic Review*, **65**, 314–325.

BRIANZONI, S., COPPIER, R. & MICHETTI, E. (2019) Evolutionary effects of non-compliant behavior in public procurement. *Structural Change and Economic Dynamics*, **51**, 106–118.

CADOT, O. (1987) Corruption as a Gamble. *Journal of Public Economics*, **33**, 223–244.

CERQUETI, R. & COPPIER, R. (2016) A game theoretical analysis of the impact of income inequality and ethnic diversity on fiscal corruption. *Annals of Operations Research*, **243**, 71–87.

CERQUETI, R. & COPPIER, R. (2018) A game for exploring political and bureaucratic corruption. *IMA Journal of Management Mathematics*, **29**, 151–173.

EHRLICH, I. (1973) The deterrent effect of capital punishment: A question of life and death. *National Bureau of Economic Research Cambridge, Mass., USA*.

FREEMAN, R. (1996) Why do so many young American men commit crimes and what might we do about it? *Journal of Economic perspectives*, **10**, 25–42.

GINTIS, H. (2000) *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. New Jersey: Princeton university press.

HEYES, A. & KAPUR, S. (2008) An economic model of whistle-blower policy. *The Journal of Law, Economics, & Organization*, **25**, 157–182.

HOFBAUER, J. & WEIBULL, J. W. (1996) Evolutionary selection against dominated strategies. *J. Econom. Theory*, **71**, 558–573.

LUI, F. (1986) A dynamic model of corruption deterrence. *Journal of public Economics*, **31**, 215–236.

MURPHY, K., SHLEIFER, A. & VISHNY, R. (1991) The allocation of talent: Implications for growth. *The quarterly journal of economics*, **106**, 503–530.

NEAR, J. & MICELI, M. (2016) After the wrongdoing: What managers should know about whistleblowing. *Business Horizons*, **59**, 105–114.

POLINSKY, A. M. & SHAVELL, S. (1979) The optimal tradeoff between the probability and magnitude of fines. *The American Economic Review*, **69**, 880–891.

ROSE-ACKERMAN, S. (1975) The economics of corruption. *Journal of public economics*, **4**, 187–203.

SAH, R. & STIGLITZ, J. (1988) Committees, hierarchies and polyarchies. *The Economic Journal*, **98**, 451–470.

SØREIDE, T. (2008) Beaten by bribery: why not blow the whistle? *Journal of Institutional and Theoretical Economics JITE*, **164**, 407–428.

TAYLOR, P. D. & JONKER, L. B. (1978) Evolutionary stable strategies and game dynamics. *Math. Biosci.*, **40**, 145–156.

TIROLE, J. (1996) A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, **63**, 1–22.

VEGA-REDONDO, F. (1996) *Evolution, games, and economic behaviour*. New York: Oxford University Press.

VILLENA, M. G. & VILLENA, M. J. (2004) Evolutionary game theory and Thorstein Veblen's evolutionary economics: is EGT Veblenian? *Journal of Economic Issues*, **38**, 585–610.