Received: 11 March 2024 | Accepted: 11 August 2024

DOI: 10.1111/jpet.12712

ORIGINAL ARTICLE

JOURNAL OF PUBLIC ECONOMIC THEORY WILEY

Corruption dynamics and political instability

Maurizio Lisciandra[1] | **Antonio Miralles Asensio**[2] | **Fabio Monteforte**[2]

[1]Department of Law, Economics, and Communication, LUMSA, Palermo, Italy

[2]Department of Economics, University of Messina, Messina, Italy

Correspondence
Antonio Miralles Asensio, Department of Economics, University of Messina, Via dei Verdi 75, 98122 Messina, Italy.
Email: antonio.mirallesasensio@unime.it

Funding information
Ministero dell'Università e della Ricerca and European Union - Next Generation EU, Grant/Award Number: 2022HLPMKN - J53D23004650006

Abstract

This paper introduces a model where a briber designs a bribing schedule targeted at the governing party within a bipartisan system to secure favorable treatment. Detected corruption increases voters' resentment, and in turn, the risk of political turnover—raising the minimum acceptable bribe. Periods without corruption mitigate such a risk. Should the briber deem bribing unprofitable for sufficiently high levels of resentment, resentment converges to a steady state in finite time. Conversely, if the briber perceives bribing as profitable regardless of the resentment level, the dynamics may result in continuous bribing and an unbounded increase in resentment (exploding dynamics). The model underscores the complexity of addressing corruption, emphasizing the need to balance reducing corruption with preventing excessive political instability. Societal forgiveness and sensitivity significantly shape corruption dynamics and public resentment. While forgiveness reduces long-run resentment, it concurrently exacerbates long-term corruption and, on balance, may have a detrimental effect on long-term welfare. Sensitivity has no long-run effect on resentment, while it reduces both corruption activity and political instability in the long term. Finally, exogenous political instability exacerbates corruption, resentment, and the risk of exploding dynamics.

KEYWORDS
anticorruption strategies, bribery dynamics, governance stability, political corruption and favoritism, voter's resentment

1 | INTRODUCTION

Despite extensive academic inquiry into political corruption, a significant gap remains in our understanding of its dynamics, especially concerning the evolving incentives for both sides of the corruption exchange as public resentment emerges in response to uncovered corruption scandals. This situation poses two critical questions: How does corruption affect political stability over time? And what role does public resentment against corruption play in shaping the incentives of both prospective bribers and bribees? Aiming to bridge this gap, this paper constructs a continuous-time theoretical model capturing the dynamic interactions among political parties, a briber, and the electorate.

More in detail, the literature highlights empirical evidence and stylized facts emphasizing the impact of corruption on public resentment on one side, and on electoral turnover on the other. For instance, Anderson and Tverdova (2003), utilizing survey data from a diverse set of countries, found that corruption undermines trust in civil servants and deteriorates the electorate's perception of the political system. This finding has been reinforced by a broad array of studies, each illuminating a specific aspect of this relationship, yet collectively affirming a general consensus on the adverse correlation between corruption and trust in political institutions.[1] More recently, corruption has been found to compromise confidence in the integrity of the entire electoral process (Cantú & García-Ponce, 2015). This is connected to the well-documented stylized fact in the empirical literature about the detrimental effect of corruption exposure on the reelection chances of incumbents (Chong, 2015; Ferraz & Finan, 2008; Klašnja & Tucker, 2013; Kostadinova, 2009; Peters & Welch, 1980). As a result, the accumulation of resentment heightens the risks associated with corruption, potentially to the extent that, if corruption becomes widespread, it could jeopardize the entire political system. Thus, it raises the question of what the optimal dynamics of bribing are and whether public resentment will ultimately reach a steady state.

In this paper, we respond to the questions outlined above by developing a theoretical model to characterize long-term corruption patterns and their electoral consequences in terms of the political system's overall stability. The model integrates classical political economy elements, including potential turnovers within a bipartisan framework and a briber seeking to co-opt the ruling party for rent extraction. It defines two potential states for a political party: in power or in opposition, and delineates two strategies: to accept or reject bribes. The briber functions as a principal within this framework, enabling her to secure the ruling party's compliance in exchange for the minimum acceptable bribe. This approach simplifies the analysis by excluding alternative corruption arrangements, thereby assigning political parties a passive role in either accepting or declining bribes.[2]

[1]Zhang and Kim (2018) provide a comprehensive review and confirm in their findings that trust in government is negatively impacted by corruption convictions.
[2]For a recent theoretical exploration of the search and matching aspects of political corruption, see Lisciandra et al. (2024).

Voters' sentiment is summarized by their resentment towards the recurrent exposure of corruption events. This resentment fluctuates with the frequency of detected corruption cases, amplifying the risk of electoral turnover and thereby increasing the cost for a briber to influence the governing party. In this context, we explore the long-term dynamics of corruption-induced resentment, which could be considered by a benevolent constitutional designer as a detrimental factor in societal welfare.

Our model is streamlined yet robust, enabling us to predict several outcomes. Two principal scenarios emerge based on the baseline parameters. In the first scenario, optimal bribing strategies lead to rapid convergence towards a stable state of resentment if the initial resentment level does not exceed a certain threshold. In the second scenario, where initial resentment exceeds this threshold, bribing yields positive immediate returns for the briber, initiating a pattern of continuous increase in societal resentment and resulting in explosive dynamics. This phenomenon is attributed to the fact that, beyond a certain level of electorate resentment, any further increase has a negligible impact on the minimum acceptable bribe, since this is contingent upon the turnover probability, inherently capped at 1. Under these conditions, corruption becomes endemic, driving public resentment to levels that threaten the political system's stability. A constitutional designer aiming to mitigate long-term electorate dissatisfaction must employ all available means to prevent this outcome. Consequently, we introduce a long-term welfare evaluation that considers political instability—represented by the turnover hazard rate–and corruption as detrimental factors. This evaluation also examines the influence of various cultural attitudes towards corruption and other variables directly affecting the political system and corruption, such as the system's inherent exogenous instability or deterrence policies.

At its core, the dynamics within the model are driven by the interplay between incentives for corruption and public resentment. This resentment, in turn, influences the reelection chances of incumbents and the consequent minimum acceptable bribe, thereby establishing a revised level of incentives for corruption. In support of this dynamic mechanism, Morris and Klesner (2010) present evidence of significant mutual causality between perceptions of corruption and trust in political institutions. Moreover, our model offers valuable insights into the influence of societal forgiveness and vigilance on the dynamics of corruption and its implications for political instability. Thus, our work connects to the broader discourse on how cultural traits impact corruption (Pellegrini & Gerlagh, 2008).

We further explore the model through various extensions, demonstrating how minor modifications can align it with empirical observations on the so-called "corruption fatigue" (De Vries & Solaz, 2017) and the nonlinear impact of exposed corruption on electoral punishment relative to the magnitude of bribery rents (Brollo et al., 2013). Additionally, although our model focuses on long-run states rather than cyclic patterns, our appendix demonstrates that an extension, where the availability of bribe rent opportunities is flexible and endogenous, could reveal cyclic behavior similar to the dynamics described in Feitchinger and Wirl (1994).[3]

We acknowledge certain limitations within the proposed model. Notably, its continuous-time framework efficiently captures the impacts of uncertain events like corruption scandals or electoral turnover within a deterministic bribing path. Yet, this simplicity may overlook

[3]Refer also to Rinaldi et al. (1998). More recently, Lee et al. (2019) adopted an evolutionary game theory framework, integrating strategies that combine altruism/egoism with optimism/prudence, resulting in a cyclical pattern of honest and corrupt behaviors among umpires.

nuanced interactions between corruption and political election cycles. For example, Figueroa (2020), examining an exposed corruption network in Argentina, observed bribes peaking 2 weeks before elections, significantly higher than those 2 weeks postelection. Similarly, Cooper (2021) reports analogous findings for African elections in democratic regimes.

While our model outlines two long-term outcomes—either stabilizing at a finite resentment level or escalating to unbounded growth—it does not accommodate both simultaneously. Given the model's parameters, only one outcome emerges as optimal. This is distinct from Dawid and Feichtinger (1996), who delve into an optimal corruption intensity path for bureaucrats, influenced by potential reputational damage, suggesting paths converge to extremes where corruption is either pervasive or nonexistent, akin to "good" and "bad" equilibria in cooperation games.

Our framework sheds light on scenarios where political elites perpetuate corruption, remaining indifferent to escalating public discontent. However, it does not directly address the repercussions of growing electorate frustration, such as the potential emergence of populist or extremist factions capable of disrupting the established political order. Farzanegan and Witthuhn (2017) demonstrate that corruption notably undermines political stability, especially when the youth demographic surpasses a certain proportion. Divergently, Gaspar et al. (2021) highlight that when a majority party corrupts the opposition, it fosters extremism among minority voters, leading to the rise of candidates less susceptible to corruption, despite potentially extreme ideologies.[4]

The paper is organized as follows: Section 2 presents the model framework. Sections 3 and 4 are dedicated to determining the optimal bribe when the briber decides to be active and formulating the optimal bribing schedule, respectively. Comparative statics and their welfare implications are examined in Sections 5 and 6. Section 7 explores two possible extensions of the model, offering insights into their potential impacts. The concluding remarks, policy implications, and directions for future research are provided in Section 8. For thoroughness, an appendix includes proofs of key results that underpin the paper's main arguments.

2 | THE MODEL

In this section, we develop a continuous-time model based on a structured sequence of decisions and events occurring within infinitesimally small time intervals, denoted as Δ. Central to our model is the briber, who represents a member of the elite, exerting corrupt influence within the political system. This briber assumes the role of a principal, singularly responsible for designing and executing the bribery schedule over time. This assumption is pivotal as it simplifies the model, enabling us to concentrate on the mechanisms through which an unchallenged elite can manipulate the political system.

In each interval $[t, t + \Delta]$, the briber accrues a rent of $r\Delta$, proportional to the length of the agreement Δ, through an illegal but favorable decision from the governing party. At time t, the briber decides whether to bribe and, if so, determines the bribe amount, denoted as $b(t)\Delta$, to be offered to the party in office, in pursuit of the aforementioned rent. This bribe is paid in advance. The ruling party, upon being offered a bribe, decides whether to accept it. No changes or new agreements can be implemented before the end of the interval, and all involved agents

[4]For a discussion on the effects of minority party capture by bribers and its impact on political turnover, see Giannoccolo and Lisciandra (2019).

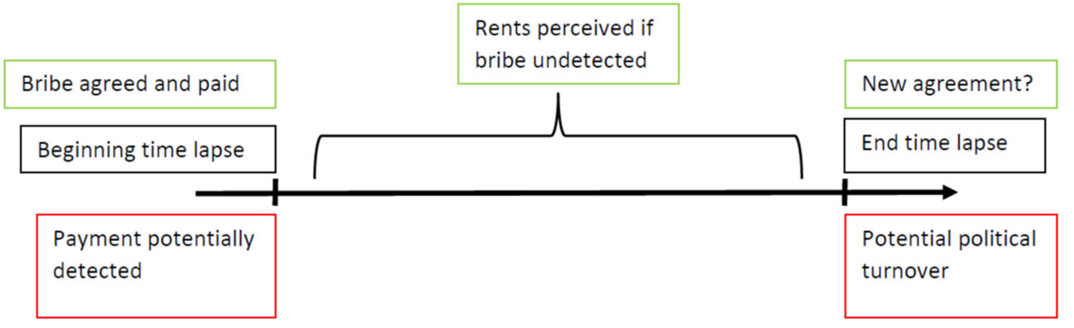

FIGURE 1 Timing of events. Illustrating the timing of events in each infinitesimal time lapse.

are compelled to adhere to the agreement until the interval's conclusion. The timing of these events is depicted in Figure 1.

We conceptualize the judiciary/surveillance system's intervention in our model as a probabilistic event. At any given time t, there exists a probability π that a bribe is discovered while it is being paid. Upon detection, both the ruling party and the briber are subjected to immediate penalties. These consequences are twofold. First, neither party realizes the intended bribe or rent; for instance, an exposed bribe is seized by the judiciary, thereby constituting a sunk cost for the briber. Second, both parties are obligated to pay a penalty, quantified as a proportion s of the intended rents.

The political system in our model is bipartisan. At the conclusion of each time interval $[t, t + \Delta]$, there is a probability that the competing party may succeed the governing party.[5] If no corruption activity occurs or is uncovered, this probability is $\theta\Delta$.[6] However, if corruption is both present and detected, the probability increases to $(\theta + \lambda(t))\Delta$, where $\lambda(t)$ represents the level of public resentment against corrupt political systems at time t. Given that Δ is infinitesimally small, it is not necessary to impose bounds on the values of θ and $\lambda(t)$ for the purposes of maintaining proper probabilities.

The dynamics of $\lambda(t)$ are as follows: $\frac{d \log \lambda(t)}{dt} = \omega > 0$ when corruption is detected in the time interval immediately preceding t, and $\frac{d \log \lambda(t)}{dt} = -\phi < 0$ when corruption is not detected.[7] Hence, public resentment accumulates against persistently corrupt political systems, while it diminishes in the absence of detected corruption. Let $\sigma = \pi\omega - (1 - \pi)\phi \gtrless 0$ denote the expected rate of increase in resentment when the ruling party is corrupt at time t. The initial level of public resentment against corruption is denoted by $\lambda(0) = \lambda_0 > 0$.

Being in office during the interval $[t, t + \Delta]$ provides the ruling party with a term-of-office payoff $m\Delta > 0$. Political parties discount future payoffs at a rate $\rho > 0$. The briber discounts future payoffs at a rate $\tilde{\rho} > 0$.

We aim to solve for the optimal bribing schedule from the briber's standpoint, who seeks to maximize the present value of her stream of payoffs over an infinite horizon.

[5]For instance, this could occur if the governing party resigns and elections are anticipated.
[6]Note that in our framework θ is a catchall variable capturing all factors *other than corruption* that affect political instability. Thus, it may also be conceived as a proxy for inefficient ruling politicians as well as the constitutional and electoral structure of the political system. On the role of improving democratic decision-making in increasing policies effectiveness, in particular within the context of sovereign debt assistance, see Ben-Yashar et al. (2023).
[7]Refer to Section 7.1 for a discussion on a law of motion with a more flexible shape.

3 | MINIMUM ACCEPTABLE BRIBES

For clarity of exposition in this section, we omit the explicit notation of time dependence. We assume that, following the rejection of a bribe, the briber commits to a no-bribe policy henceforth. This commitment strategy is deemed most profitable for the briber. It minimizes the continuation payoff for the political parties should they decline a bribe, thereby reducing the minimum bribe they would be willing to accept.[8]

In the lines that follow, we provide some arguments and evidence on the existence of commitment.[9] Nevertheless, it has to be noted that the existence of a commitment technology is for the purposes of simplicity of exposition. Actually, one can obtain a payoff-equivalent subgame perfect equilibrium in the briber–bribee dynamic game without commitment. One first notices that, in the absence of the briber, the ruling party obtains always the same expected present value of payoff flows, say EU, regardless of how resentful against corruption the electorate is, since no corruption scandals can happen. Then we consider a bribing strategy in the following class: whenever a bribe is offered, it will be a bribe that makes the accepting ruling party obtain a present value equal to EU, regardless of the history of previous offers and acceptance/rejection decisions. A weakly optimal strategy for the ruling party then is to always accept the bribes when offered. Moreover, the offered bribe is the best from the briber's point of view. The briber–bribee interaction is therefore represented by a repeated ultimatum game which contains a subgame perfect equilibrium that matches the flow of instant payoffs we obtain in the present model. While other equilibria are possible, we focus on the one that is most beneficial for the briber.

We believe commitment is more natural to assume if looking at the specific (*secret!*) context in which the agreement takes place. Indeed, the briber–bribee relationship is essentially based on trust and each party has strong incentives to establish a reputation of trustworthiness within that specific "market." In addition, when the briber makes an offer, he discloses himself as "one of that kind," so to insure against the possible related risks, he needs to provide a very credible threat against not accepting her offer.[10]

Alternatively, we could think about another mechanism based on trust and requiring commitment. Specifically, let us assume that, when the bribe is not accepted by the politician, the bribing attempt becomes easily detected and that, upon detection, the briber exits the market (e.g., she goes to jail) and another one jumps up with a possible new offer. However, given its previous refusal history, the politician will be regarded as not trustworthy and no offer

[8]The model could be alternatively constructed on the basis of Nash bargaining for the bribe paid per period. With Nash bargaining, the bribe paid would not be the lowest possible, as it would depend upon the size of the (exogenously given) relative bargaining power of the briber and the ruling party. Dynamics however would be qualitatively similar to what shown in the sections to follow, with the difference that the net instant payoff obtained by the briber would be a fraction $\zeta \in (0, 1)$ of what obtained in the baseline model.

[9]We thank a referee who made us reflect on the necessity of giving a solid base to the assumption of commitment.

[10]See, for example, how reputation building concerns of trustworthiness in political corruption context were indeed present during *Tangentopoli*, the largest corruption network ever discovered in Italy. Della Porta and Vannucci (2012) illustrate this mechanism through police recordings of conversations between corrupt parties, revealing that corrupt officials were careful in developing, spreading, and maintaining a reputation for being reliable. These officials accepted bribes without reporting bribery attempts to the police and reciprocated bribes with good performance, otherwise hardly performing at all.

will actually be placed. Thus, for any bribery to be sustained over the long run, commitment is needed.[11]

We now go back to the model. Let X and Y represent the ruling and competing political parties at time t, respectively. If a party is corrupted, it is denoted with the subscript c, whereas the subscript h identifies "honest" behavior, implying either the rejection of the bribe by the ruling party or the adoption of the briber's no-bribe policy. Note that these are not intrinsic characteristics but rather outcomes of the decisions made by the parties at the time.

For an honest party, the instantaneous probability of political turnover is independent of the decision moment t, being solely $\theta\Delta$. Consequently, the present value of expected utility—including both current and future utility—is temporally constant for both parties X and Y. With an infinitesimally small time lapse Δ, the ruling party X confronts the following characterization of this constant present value:

$$EU_{X_h} = m\Delta + \mathrm{e}^{-\rho\Delta}[\theta\Delta \cdot EU_{Y_h} + (1 - \theta\Delta)EU_{X_h}]. \tag{1}$$

Utilizing the approximation $\mathrm{e}^{-\rho\Delta} \approx 1 - \rho\Delta$, and taking the limit as $\Delta \to 0$, we derive the so-called "asset equation":

$$\rho EU_{X_h} = m - \theta[EU_{X_h} - EU_{Y_h}]. \tag{2}$$

Similarly, we derive the asset equation for the honest opposition party Y:

$$\rho EU_{Y_h} = \theta[EU_{X_h} - EU_{Y_h}]. \tag{3}$$

Solving this system, we obtain

$$EU_{X_h} = \frac{\rho + \theta}{\rho[\rho + 2\theta]} m, \tag{4}$$

$$EU_{Y_h} = \frac{\theta}{\rho[\rho + 2\theta]} m. \tag{5}$$

When parties are corrupt, the expected probability of turnover $(\theta + \lambda\pi)\Delta$ varies over time as λ depends on t. Initially, it cannot be assumed that the present value of their expected payoffs remains constant over time. However, by offering the minimum bribe that makes the ruling party indifferent to accepting or rejecting it, the briber ensures that the present value of expected payoffs remains constant. Consequently, both EU_{X_c} and EU_{Y_c} become constant in t.

The present values of expected utilities under corruption are characterized by the following asset equations:

$$\rho EU_{X_c} = \tilde{m} - (\theta + \lambda\pi)[EU_{X_c} - EU_{Y_c}], \tag{6}$$

[11]This is in line with the analysis provided in Dufwenberg and Spagnolo (2015, p. 841) where "the threat of perpetual reversion to a stage game Nash equilibrium can credibly be used to sustain other equilibria where the long-run player avoids her static best response." We thank the editor for pointing us towards this analysis.

$$\rho EU_{Y_c} = (\theta + \lambda\pi)[EU_{X_c} - EU_{Y_c}], \tag{7}$$

where $\tilde{m} = m + (1 - \pi)b - \pi sr$. The solution to this system of equations is given by

$$EU_{X_c} = \frac{\rho + \theta + \lambda\pi}{\rho[\rho + 2(\theta + \lambda\pi)]}\tilde{m}, \tag{8}$$

$$EU_{Y_c} = \frac{\theta + \lambda\pi}{\rho[\rho + 2(\theta + \lambda\pi)]}\tilde{m}. \tag{9}$$

We now apply the indifference condition, $EU_{X_c} = EU_{X_h}$, yielding

$$\frac{\tilde{m}}{m} = \frac{(\rho + \theta)(\rho + 2\theta) + 2\lambda\pi(\rho + \theta)}{(\rho + \theta)(\rho + 2\theta) + \lambda\pi(\rho + 2\theta)}. \tag{10}$$

Given that $\frac{\tilde{m}}{m} = 1 + (1 - \pi)\frac{b}{m} - \pi\frac{sr}{m}$, the optimal bribe (conditional on bribing) is determined as

$$b = \frac{\pi}{1 - \pi}\left[sr + \frac{\lambda\rho m}{(\rho + \theta)(\rho + 2\theta) + \lambda\pi(\rho + 2\theta)}\right]. \tag{11}$$

It is notable that the value of the minimum acceptable bribe b increases with the extent of public resentment against corruption λ. Consequently, when offered as calculated, the bribe is invariably accepted by the ruling party. This ensures that the briber's committed no-bribe policy against rejected offers is never activated.

This leaves the instant payoff for an active briber as

$$\gamma(\lambda) = (1 - \pi)r - \pi sr - \frac{\pi}{1 - \pi}\left[sr + \frac{\lambda\rho m}{(\rho + \theta)(\rho + 2\theta) + \lambda\pi(\rho + 2\theta)}\right]. \tag{12}$$

Upon observing the characteristics of γ, we note that it decreases with an increase in the detection probability, π, the size of the related fine, s, and the term-of-office payoffs, m, while it increases with an increase in the rents from bring, r and the exogenous political instability, θ. The implied change in an active briber's instant payoff to variations in the parties' discount rate, ρ remains ambiguous. In general, the intuitive impacts of these parameters are notable. Specifically, an increase in exogenous political instability enhances the briber's instant payoff, which can be attributed to the fact that a higher θ raises the baseline probability of political turnover. This, in turn, results in a relatively smaller *increase* in the risk of turnover due to uncovered corruption, leading to a lower minimum bribe acceptable to the ruling party.

4 | THE OPTIMAL SCHEDULE OF BRIBING

We make the following notation shortcuts:

$$A = (1 - \pi)r - \pi sr - \frac{\pi}{1 - \pi}sr, \tag{13}$$

$$B = \frac{\pi\rho m}{1-\pi}, \tag{14}$$

$$c = (\rho + \theta)(\rho + 2\theta), \tag{15}$$

$$d = \pi(\rho + 2\theta). \tag{16}$$

Without loss of generality, we can assume that these four values are positive. This notation allows us to express the briber's instant payoff, when actively bribing, as a function of λ:

$$\gamma(\lambda) = A - B\frac{\lambda}{c + d\lambda}. \tag{17}$$

The briber designs her schedule of bribing activity to maximize the present value of present and future payoffs. Let $q(t)$ denote the chosen probability of being an active briber at time t. The maximization problem is formulated as

$$Q = \max_{q} \int_0^{\infty} q(t)\mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda(t))\,dt \tag{18}$$

subject to $0 \leqslant q(t) \leqslant 1$, $\frac{\dot{\lambda}}{\lambda} = -\phi + (\sigma + \phi)q$, and $\lambda(0) = \lambda_0$.

Let $\mu(t) \geqslant 0$ denote the Lagrange multiplier associated with $q(t) \geqslant 0$, $\nu(t) \geqslant 0$ denote the Lagrange multiplier associated with $1 - q(t) \geqslant 0$, and $\beta(t) \geqslant 0$ denote the shadow cost associated with the law of motion for λ. Omitting t for simplicity, we write the Hamiltonian:

$$H = q\mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda) - \beta\lambda(-\phi + (\sigma + \phi)q) + \mu q + \nu(1 - q). \tag{19}$$

Pontryagin's optimality conditions are

$$\frac{\partial H}{\partial q} = \mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda) - \beta\lambda(\sigma + \phi) + \mu - \nu = 0 \tag{20}$$

with $\mu = 0$ if $q > 0$ and $\nu = 0$ if $q < 1$, and

$$\frac{\partial H}{\partial \lambda} = q\mathrm{e}^{-\tilde{\rho}t}\gamma'(\lambda) - \beta(-\phi + (\sigma + \phi)q) = \dot{\beta}. \tag{21}$$

We define an economically intuitive expression, which will be helpful to study the dynamics of resentment:

$$g(t) := \mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda(t)) - \beta(t)\lambda(t)(\sigma + \phi). \tag{22}$$

This expression captures, on the one hand, the discounted instant payoff from bribing activity, and on the other hand, the dynamic penalty associated with the consequent increase in the state of anger. It is noteworthy that $g > 0$ implies $q = 1$ (and thus $\dot{\lambda} > 0$), whereas $g < 0$

necessitates $q = 0$ (and thus $\dot{\lambda} < 0$), while when $g = 0$, any bribery probability within the range of $q \in [0, 1]$ is deemed optimal.

By synthesizing Pontryagin's optimality conditions, we deduce that

$$\dot{g} = -\mathrm{e}^{-\tilde{\rho}t}[\tilde{\rho}\gamma(\lambda) + \phi\lambda\gamma'(\lambda)]. \tag{23}$$

We then define

$$k(\lambda) := \tilde{\rho}\gamma(\lambda) + \phi\lambda\gamma'(\lambda). \tag{24}$$

This allows us to simplify the expression as follows:

$$\dot{g} = -\mathrm{e}^{-\tilde{\rho}t}k(\lambda(t)). \tag{25}$$

Hence, it is evident that the sign of $\dot{g}$ is solely determined by the state of resentment, λ, as expressed through the function k, independently of other variables.

We now introduce a useful change of variable: $z := \frac{1}{c + d\lambda}$, which implies $cz = \frac{c}{c + d\lambda} \in (0, 1)$. This relation signifies that as $\lambda \to \infty$, $cz \to 0$, and as $\lambda \to 0$, $cz \to 1$. Given the definition of z, we further derive $\lambda = \frac{1}{dz} - \frac{c}{d}$. Recall finally that $\gamma(\lambda) = A - B\frac{\lambda}{c + d\lambda}$ and $\gamma'(\lambda) = \frac{-Bc}{(c + d\lambda)^2}$. With this change of variable, the derivative $\gamma'(\lambda)$ can be re-expressed in terms of cz as $\gamma'(\lambda) = -Bc(z^2)$.

When k is expressed as a function of cz, we observe a quadratic U-shaped form for k. This form is derived as follows:

$$k = \tilde{\rho}A + \frac{B}{d}\left(c - \frac{1}{z}\right)(\tilde{\rho}z + c\phi z^2) = \tilde{\rho}A + \frac{B}{d}[-\tilde{\rho} + (\tilde{\rho} - \phi)cz + \phi(cz)^2]. \tag{26}$$

Considering the defined relationship between cz and λ, the U-shaped characteristic of k dictates the following: should k exhibit an increasing trend at a given value of λ, it is then expected to continue increasing for all subsequent values $\lambda' > \lambda$. Conversely, should k demonstrate a decreasing trend at a specific λ, it will be decreasing for all preceding values $\lambda' < \lambda$. We finally observe that $\lim_{cz \to 0} k = \lim_{\lambda \to \infty} k = \tilde{\rho}\left(A - \frac{B}{d}\right)$, and that $\lim_{cz \to 1} k = \lim_{\lambda \to 0} k = \tilde{\rho}A > 0$.

There are two important scenarios to consider, as illustrated in Figure 2. In one scenario, the intercept on the vertical axis $A - \frac{B}{d} > 0$, signifies that, irrespective of the level of resentment state, the briber can always realize positive instant payoffs. In this scenario, it is possible to identify either two distinct values, $\lambda_1^* < \lambda_2^*$, or a single value, λ^*, for which $k = 0$ (and thus $\dot{g} = 0$). There may also be instances where no solution exists.[12] Conversely, in the other scenario, where $A - \frac{B}{d} < 0$, a unique λ^* exists. The dynamics differ between these scenarios.

One might ponder whether, upon reaching the candidate steady state λ^*, it can be sustained with

[12]If $A - \frac{B}{d} > 0$, the condition $\phi > \rho$ is necessary for the existence of at least one solution.

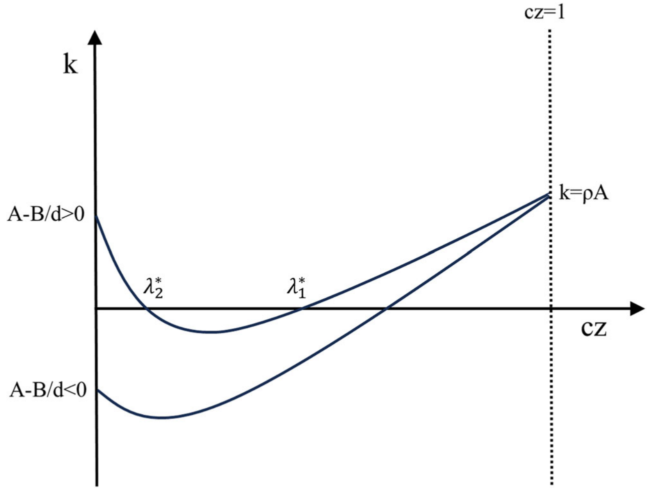

FIGURE 2 Scenorios for public resentment equilibria. Instances of two distinct scenarios are significant for understanding the dynamics of corruption and resentment. When $A - \frac{B}{d} > 0$, there may be two values of $\lambda = \frac{c}{d}\left(\frac{1}{cz} - 1\right)$ that result in $k = 0$. Conversely, if $A - \frac{B}{d} < 0$, there is always a single such value.

$$q^* = \frac{\phi}{\sigma + \phi}, \tag{27}$$

the sole probability that maintains $\dot{\lambda} = 0$. For such q^* to be optimal, it necessitates $g = 0$. Intrinsically, by the definition of λ^*, we ascertain $k = 0$ and consequently $\dot{g} = 0$, which implies g remains constant over time; however, the specific level of g remains undetermined. The subsequent lemma shows that $g = 0$ over a time interval if and only if $k = 0$, thereby affirming λ^* as a viable steady state.

Lemma 1. *Let $g(t) = 0$ for all $t \in (T, T')$. Then, $\lambda(t) = \lambda^*$ throughout the interval, where λ^* satisfies*

$$k := \tilde{\rho}\gamma(\lambda^*) + \phi\lambda^*\gamma'(\lambda^*) = 0. \tag{28}$$

The proof is provided in the appendix.

From the discussion above, the following remark follows:

Remark 1 (Weak monotonicity). Along the optimal bribing activity path, λ is weakly monotonic over time. Whenever λ increases, we have $q = 1$; when λ decreases, $q = 0$; and, when λ remains constant, $q = q^*$.

The proof is provided in the appendix. In other words, given the U-shaped characteristics of k, when the briber finds it optimal to be always active, public resentment increases, since there will be an uninterrputed sequence of bribing episodes, while it decreases when the briber finds it optimal not to be active, and it will be constant when the briber will not always be active.

This observation identifies four viable candidates for the optimal path of bribing activity:

1. $q = 0$ at all times;
2. $q = 0$ until $\lambda = \lambda^*$, at which point $k(\lambda^*) = 0$, and $q = q^*$ thereafter;
3. $q = 1$ until $\lambda = \lambda^*$, at which point $k(\lambda^*) = 0$, and $q = q^*$ thereafter;
4. $q = 1$ at all times.

However, the optimal path will depend on the initial level of public resentment, λ_0. In what follows, we will thus characterize the optimal strategies of bribing activities, by comparing the above four options according to the initial level of public resentment. This comparison will then allow us to describe the optimal bribing dynamics and the related paths for public resentment.

The second and third options are mutually exclusive unless $\lambda_0 = \lambda^*$, in which case the two options become equivalent. If $\lambda_0 < \lambda^*$, then the second option is precluded; conversely, if $\lambda_0 > \lambda^*$, the third option is precluded.

To compare the third and fourth options, it is necessary to evaluate the present value of payoffs when the briber maintains resentment constant at λ^* against the present value of payoffs when the briber, facing such a level of resentment, opts to initiate active bribing thereafter. Thus, we start the comparison when λ^* is reached. The difference is as follows:

$$\begin{aligned} N &:= \frac{\phi}{\tilde{\rho}(\sigma+\phi)}\gamma(\lambda^*) - \int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda^*\mathrm{e}^{\sigma t})\,dt \\ &= -\frac{\phi}{\tilde{\rho}(\sigma+\phi)}\int_0^\infty \frac{d\left[\mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda^*\mathrm{e}^{\sigma t})\right]}{dt}dt - \int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda^*\mathrm{e}^{\sigma t})\,dt \\ &= -\frac{\sigma}{\sigma+\phi}\left(\int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda^*\mathrm{e}^{\sigma t})\,dt + \frac{\phi}{\tilde{\rho}}\int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma'(\lambda^*\mathrm{e}^{\sigma t})\lambda^*\mathrm{e}^{\sigma t}\,dt\right) \\ &= -\frac{\sigma}{\tilde{\rho}(\sigma+\phi)}\int_0^\infty \mathrm{e}^{-\tilde{\rho}t}k(\lambda^*\mathrm{e}^{\sigma t})\,dt, \end{aligned} \tag{29}$$

where the initial equality is derived from $\gamma(\lambda^*) = -\left[\mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda^*\mathrm{e}^{\sigma t})\right]_0^\infty$, and the subsequent equality results from the rearrangement following the differentiation of $\gamma(\lambda^*)$.

When comparing alternatives two and four, we need to start from λ_0. We must compare the present value of payoffs until λ decreases to λ^*, at which point that level is maintained constant, with the present value of payoffs resulting from being active starting from λ_0. Let $T(\lambda_0, \lambda^*, \phi)$ denote the time required for λ to reach the value λ^*. The difference between the present values is:

$$\begin{aligned} M(\lambda_0) &:= \mathrm{e}^{-\tilde{\rho}T(\lambda_0,\lambda^*,\phi)}\frac{\phi}{\tilde{\rho}(\sigma+\phi)}\gamma(\lambda^*) - \int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda_0\mathrm{e}^{\sigma t})\,dt \\ &= (\lambda^*/\lambda_0)^{\tilde{\rho}/\phi}\frac{\phi}{\tilde{\rho}(\sigma+\phi)}\gamma(\lambda^*) - \int_{\lambda_0}^\infty (\lambda/\lambda_0)^{-\tilde{\rho}/\sigma}\frac{\gamma(\lambda)}{\sigma\lambda}d\lambda, \end{aligned} \tag{30}$$

where the latter equality originates from $\lambda^* = \lambda_0\mathrm{e}^{-\phi T(\lambda_0,\lambda^*,\phi)}$, on the one hand, and from a change of variable $\lambda = \lambda_0\mathrm{e}^{\sigma t}$ within the integral, on the other hand. Of course, $N = M(\lambda^*)$.

Observe that $\lim_{\lambda_0\to\infty} M(\lambda_0) = \lim_{\lambda_0\to\infty} \lambda_0^{-\tilde{\rho}/\phi}M(\lambda_0) = 0$. Through some algebraic manipulation, it can be shown that $\frac{d\left[\lambda_0^{-\tilde{\rho}/\phi}M\right]}{d\lambda_0} < 0$ if and only if $(\lambda^*)^{\tilde{\rho}/\phi}\gamma(\lambda^*) > \lambda_0^{\tilde{\rho}/\phi}\gamma(\lambda_0)$.

Consider the function $\Gamma(\lambda) := \lambda^{\tilde{\rho}/\phi}\gamma(\lambda)$. Its derivative shares the same sign as $k(\lambda)$. If $A - B/d > 0$ and there is only one root λ^* for k, Γ is always increasing except at the inflection point λ^*. If $A - B/d > 0$ and there are two roots $\lambda_1^* < \lambda_2^*$, then λ_1^* marks a local maximum for Γ, while λ_2^* identifies a local minimum. Given that $\lim_{\lambda\to\infty}\Gamma(\lambda) = \infty$, there exists a $\bar{\lambda} > \lambda_2^*$ such

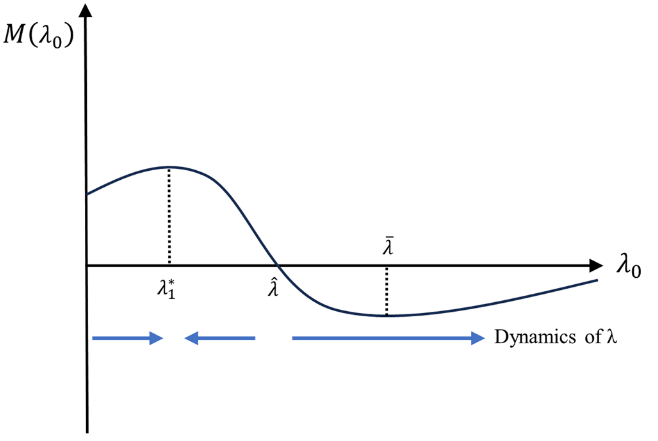

FIGURE 3 Public resentment threshold for exploding dynamics. Plot of M against λ_0 for $A - B/d > 0$ and $N > 0$. Initial states above $\hat{\lambda}$ result in explosive dynamics, while those below $\hat{\lambda}$ lead to fast convergence towards λ_1^*.

that $\Gamma(\bar{\lambda}) = \Gamma(\lambda_1^*)$. Conversely, in the scenario where $A - B/d < 0$, $k(\cdot)$ is positive to the left of the unique root λ^* and negative to its right, indicating that the unique root λ^* establishes a global maximum for Γ.

Suppose $A - B/d > 0$ and assume there is only one root, λ^*. In such circumstances, $k(\lambda) > 0$ for every $\lambda > \lambda^*$. Consequently, this implies $N < 0$. Given that $M(\lambda_0)$ is initially decreasing and subsequently increasing, and considering it converges to zero, it follows that $M(\lambda_0) < 0$ for all $\lambda_0 > \lambda^*$. Hence, alternative 4 emerges as optimal, regardless of the initial value λ_0.

Suppose $A - B/d > 0$ and assume that there are two roots, $\lambda_1^* < \lambda_2^*$. The reasoning in the preceding paragraph rules out λ_2^* as the point at which to maintain the resentment state constant.[13] Consequently, we focus on alternatives 2 and 3, where the resentment state reaches a steady state at λ_1^*.

First, let us compute N for the steady-state value λ_1^*. If $N < 0$, we deduce that alternative 4 is preferred over all others, irrespective of the initial value λ_0.[14] Now consider $N > 0$; thus, if $\lambda_0 \leqslant \lambda_1^*$, alternative 3 is favored over alternative 4. Given $\lambda_0^{-\tilde{\rho}/\phi} M(\lambda_0)$ is decreasing for $\lambda_0 \in (\lambda_1^*, \bar{\lambda})$ and increases thereafter, with $\lambda_0^{-\tilde{\rho}/\phi} M(\lambda_0)$ ultimately converging to zero, and considering $N = M(\lambda_1^*) > 0$, there exists a unique value $\hat{\lambda}$ at which M changes sign from positive to negative.[15] Consequently, for $\lambda_0 \in (\lambda_1^*, \hat{\lambda})$, alternative 2 is superior to alternative 4, while for $\lambda_0 > \hat{\lambda}$, alternative 4 is deemed optimal. This dynamics is shown in Figure 3.

In the final scenario, where $A - B/d < 0$, we find that there is one root λ^* for k, which also sets a global maximum for Γ. Consequently, $\lambda_0^{-\tilde{\rho}/\phi} M(\lambda_0)$ decreases for all $\lambda_0 > \lambda^*$, and since

[13]For $\lambda_0 < \lambda_2^*$, $N\left(\lambda_2^*\right) < 0$. Furthermore, $M(\lambda_0)$ is decreasing for $\lambda_1^* < \lambda_0 < \lambda_2^*$ and increasing for $\lambda_0 > \lambda_2^*$, and given that it converges to zero, it follows that $M(\lambda_0) < 0$ for all $\lambda_0 > \lambda_2^*$.

[14]If $N < 0$ and $\lambda_0 \leqslant \lambda_1^*$, then alternative 4 is preferred to alternative 3. Given $N = M\left(\lambda_1^*\right) < 0$ and since $M\left(\lambda_1^*\right)$ decreases after λ_1^* and then increases but converges to zero, it follows that for $\lambda_0 > \lambda_1^* M(\lambda_0) < 0$ and thus alternative 4 remains preferred to alternative 2.

[15]This indicates $M'(\hat{\lambda}) < 0$, a crucial point for comparative statics analysis.

$\lambda_0^{-\tilde{\rho}/\phi} M(\lambda_0)$ eventually converges to zero, this implies that $M(\lambda_0) > 0$ for every $\lambda_0 > \lambda^*$. Given this result, with $N = M(\lambda^*) > 0$, we can exclude alternative 4 as a viable option, regardless of the initial value of λ_0.

We summarize the main results of all this digression into Proposition 1. We say that bribing dynamics show *fast convergence* to some resentment value λ^*: (a) if $\lambda_0 \leqslant \lambda^*$ and $q = 1$ until $\lambda = \lambda^*$, while $q = q^*$ thereafter; or (b) if $\lambda_0 > \lambda^*$ and $q = 0$ until $\lambda = \lambda^*$, while $q = q^*$ thereafter. We say that bribing dynamics are *explosive* if $q = 1$ at every moment.

Proposition 1. *Optimal bribing dynamics are as follows:*

1. *If $A - B/d < 0$, then there is fast convergence to the unique root of k.*
2. *If $A - B/d > 0$, and k has one or no roots, then dynamics are explosive.*
3. *If $A - B/d > 0$, and k has two roots $\lambda_1^* < \lambda_2^*$, and $\int_0^\infty e^{-\tilde{\rho}t} k\left(\lambda_1^* e^{\sigma t}\right) dt > 0$, then dynamics are explosive.*
4. *If $A - B/d > 0$, and k has two roots $\lambda_1^* < \lambda_2^*$, and $\int_0^\infty e^{-\tilde{\rho}t} k\left(\lambda_1^* e^{\sigma t}\right) dt < 0$, then there is $\hat{\lambda} > \lambda_1^*$ such that, for $\lambda_0 < \hat{\lambda}$ we have fast convergence to λ_1^*, and for $\lambda_0 > \hat{\lambda}$ we have explosive dynamics.*

In other words, if, for the range of parameter values, the briber can realize positive instant payoffs even when resentment is very high ($A - B/d > 0$), dynamics will be always explosive (points 2 and 3 of Proposition 1) unless there is a range of resentment values for which the present value of payoffs when the briber keeps resentment constant is higher than the respective present value when he opts to being always active. In this latter case, there will be a threshold value of resentment triggering explosive dynamics (point 4 of Proposition 1). On the other hand, if the range of parameter values does not imply positive payoffs when resentment is very high ($A - B/d < 0$), dynamics will quickly converge to a steady state where resentment is kept constant and bribing happens with some positive probability, but not with certainty (point 1 of Proposition 1).

5 | COMPARATIVE STATICS

In this section, we explore how changes in the model's parameters influence the equilibrium state to provide a comprehensive understanding of the dynamics affecting corruption and societal resentment levels. An initial, straightforward result pertains to the comparative statics of the long-run corruption activity level in steady state, q^*, as detailed in the following lemma:

Lemma 2. *The steady-state level of corruption, q^*, depends positively on societal forgiveness, ϕ, and negatively on the sensitivity to corruption scandals, ω, and the detection probability, π.*

The positive dependence of q^* on ϕ (societal forgiveness or forgetfulness) can be attributed to the reduced social and legal repercussions for corrupt behavior in a society that is more forgiving or forgetful, effectively lowering the barriers to corruption. In contrast, the negative dependence of q^* on ω (the sensitivity of the population to corruption scandals) is due to the heightened risks and costs of corruption in a society that is more vigilant and intolerant of such actions, which bolsters the efficacy of anticorruption measures and diminishes the steady-state level of corruption.

Assuming now the existence of a resentment steady state λ^*, this steady state is defined by the conditions $k(\lambda^*) = \tilde{\rho}\gamma(\lambda^*) + \phi\lambda^*\gamma'(\lambda^*) = 0$ and $k'(\lambda^*) < 0$. The latter condition is utilized to identify the minimum root when two are present: consequently, in subsequent discussions, when k exhibits two roots $\lambda_1^* < \lambda_2^*$, λ^* denotes λ_1^*.[16] This selection criterion has been elaborated upon in Section 4.

It is worthwhile to analyze how the steady resentment state λ^* reacts to changes in parameters. Questions of interest include: How does λ^* respond to increased effectiveness in uncovering and punishing illegal activities? How does exogenous political instability affect the long-run resentment state λ^*? In what manner do cultural trends of refusal and forgiveness of committed corruption activities influence λ^*? Lastly, what role does the discount rate of the bribers and of the political parties play in determining the value of λ^*?

Lemma 3. *Higher briber's discount rate, $\tilde{\rho}$, rents from bribing r, or exogenous political instability, θ increase the steady resentment state, λ^*, while higher societal forgiveness, ϕ, pecuniary fines s, detection probability, π, or term-of-office payoffs, m decrease the steady resentment state λ^*. The effect of a higher parties' discount rate ρ on the steady state of resentment λ^* might be either an increase or a decrease. Finally, the public sensitivity to corruption scandals, ω does not influence the value of the steady state of resentment, λ^*.*

The proof is provided in the appendix.

Two observations worth pointing out. First, the increase in resentment levels following uncovered corruption scandals, denoted by ω, which represents the population's sensitivity to such scandals, surprisingly does not influence the long-run resentment level, λ^*. This suggests that, despite initial reactions to corruption exposure, the enduring sentiment of resentment within the populace remains unchanged in the long term. However, as one might anticipate, this detachment does not extend to the steady-state level of corruption, q^*, which is adversely affected by ω in a predictable manner: the sensitivity against corruption negatively impacts the level of corruption.

The second observation concerns the impact of exogenous, noncorruption-related political instability, symbolized by θ, on long-run resentment levels. An increase in θ increases λ^*, indicating that factors outside of corruption also play a significant role in shaping public resentment. A plausible explanation for this phenomenon is that higher levels of inherent political instability increase the ruling party incentives to gain as much as possible from its government spell. The opportunity cost of accepting a bribe—the future returns of being in office—is expected lower. Consequently, this leads to a reduction in the threshold of bribe acceptance, effectively lowering the minimum bribe that is deemed acceptable in such environments.[17]

Among the elements contained in θ's catchall box, one could think of the abundance of poorly performing politicians due to the lack of both good formation and managerial skills. These politicians tend to take suboptimal policy decisions. Frequent blunders lead to higher

[16]Please note that the subsequent results also apply to the other unique stable equilibrium, λ^*, when $A - \frac{B}{d} < 0$.

[17]We thank a referee for this intuitive explanation.

electoral volatility (higher θ). Now, a question usually tackled is whether corruption causes more or less inefficient performance. Whether or not this is true, it is a question that places corruption as the cause and inefficiency as the consequence. We go instead in the opposite direction: poorly performing politicians lead to more long-term corruption and to more risk of exploding dynamics in corruption.

In the analysis presented in Section 4, we established that for nonexplosive dynamics to prevail—wherein $A - \frac{B}{d} > 0$—it is necessary for k to have two distinct roots, denoted as λ_1^* and λ_2^*, with $\lambda_1^* < \lambda_2^*$, and the initial condition $\lambda_0 < \hat{\lambda}$. Dynamics that fail to meet these criteria lead to explosive growth in resentment, culminating in pervasive corruption and extreme political instability. Such outcomes are unequivocally detrimental to long-term welfare. Therefore, understanding the comparative statics associated with this threshold becomes crucial.

We can formulate the following lemma:

Lemma 4. *Higher values of societal forgiveness, ϕ, pecuniary fines, s, or term-of-office payoffs, m lead to an increase in the threshold triggering exploding dynamics, $\hat{\lambda}$, whereas higher values of the briber's discount rate, $\tilde{\rho}$, rents from bribing, r, or exogenous political instability, θ result in a decrease in the exploding dynamics threshold, $\hat{\lambda}$. The effects of detection probability, π, the public sensitivity to corruption scandals, ω, and the parties' discount rate, ρ on the exploding dynamics threshold, $\hat{\lambda}$ are ambiguous.*

The proof is provided in the appendix. Lemma 4 provides insight into the dynamics that could lead to endemic corruption and a deep mistrust in political institutions, potentially destabilizing the political system. It suggests that conditions which increase $\hat{\lambda}$ help mitigate such outcomes by limiting the initial states that result in explosive dynamics. Within this framework, two particular observations merit discussion for their less intuitive implications at first glance.

A quicker societal forgiveness or forgetfulness (ϕ) is shown to reduce the range of initial states leading to explosive dynamics. This mechanism can serve as a potential preventative measure against the vicious cycle of corruption fostering mistrust, which then feeds back into more corruption and instability. In other words, by enabling a faster recovery from episodes of uncovered corruption, a larger ϕ lessens the accumulation of resentment, thereby decreasing the potential for situations that could escalate into widespread instability. This dynamic also provides political actors with more space to enact genuine reforms without facing immediate backlash. However, this perceived advantage is offset by the “too forgiving” trap, where an elevated ϕ might inadvertently raise steady-state corruption levels (q^*), by diminishing the consequences for corrupt actions. This scenario risks creating a milieu so permissive that political corrupt entities face minimal repercussions, thereby encouraging the proliferation of unchecked corruption. In such an environment, corrupt networks could become more entrenched and sophisticated, learning to navigate a more forgiving system to sustain their corrupt practices.

Furthermore, Lemma 4 touches on the impact of higher exogenous political instability (θ) on societal welfare, noting its propensity to increase the incentives for bribery. This increase in baseline political turnover risk diminishes the cost of additional turnover risk posed by uncovered corruption, thus lowering the minimum bribes deemed acceptable by the ruling party and fostering an environment more conducive to corruption.

These two points underscore the complex relationship between societal attitudes, political stability, and the prevalence of corruption, highlighting the delicate balance needed in addressing these issues.

6 | WELFARE ANALYSIS

As we have seen from the analysis in Section 5, voters' discontent against corruption has a twofold effect on the political environment. On the one hand, higher sensitivity to corruption can deter incentives to bribing by increasing the amount of the acceptable bribe, but, on the other hand, it also increases political instability. This calls naturally into question what are the long-run welfare implications of the model. With the above considerations in mind, we postulate a welfare evaluation of long-run values that takes into account several features. First, it is negatively affected by the long-run corruption activity probability. It is also affected by political instability, due to continuous turnovers. Instability has the drawback that any policy implementation becomes virtually impossible. Of course, complete stability, on the other extreme, cannot be good either, since bad governments need to be substituted.

In the long run, the duration of a term of office behaves as an exponential distribution with hazard rate $\theta + \pi\lambda^* q^*$. The expectation of the term's duration is the inverse of this hazard rate. We postulate this hazard rate as a measure of political instability. Along with other arguments, our welfare function is written as

$$W := W(q^*, \theta + \pi\lambda^* q^*) \tag{31}$$

with $W(1, \infty) = -\infty$. Recall that $q^* = \frac{\phi}{\sigma + \phi} = \frac{\phi}{\pi(\omega + \phi)}$, thus

$$W = W\left(\frac{\phi}{\pi(\omega + \phi)}, \theta + \frac{\phi}{\omega + \phi}\lambda^*\right) \tag{32}$$

The first argument in $W(\cdot,\cdot)$ affects welfare negatively. The second argument affects welfare negatively as well if and only if the hazard rate is high enough: a certain degree of instability in the political system is physiological or even desirable, but if it is too high then it may lead to the paralysis of the political system.

In the welfare function, we abstract from other arguments that in principle should be included, but whose impact on welfare is straightforward. For example, explosive dynamics are unambiguously very bad in our welfare evaluation. Similarly, another argument considered as implicit may include the corruption rent opportunity parameter r, which has a direct negative impact on welfare through its distortionary effect on the economy. However, in the current setup, we focus only on its indirect impact through λ^* and $\hat{\lambda}$. Finally, the proposed welfare function is influenced by policy parameters that incur implementation costs. These costs have not been considered in our analysis, and thus, from a policy perspective, should be deducted.

A welfare function that balances between corruption activity and political instability is the basis that justifies the comparative statics analysis above, and a graphical representation of its correlated indifference curves is provided in Figure 4.

In the subsequent analysis, leveraging the insights from Lemmas 2 to 4 in the comparative statics section, we provide an in-depth analysis of the welfare implications of exogenous parameters.

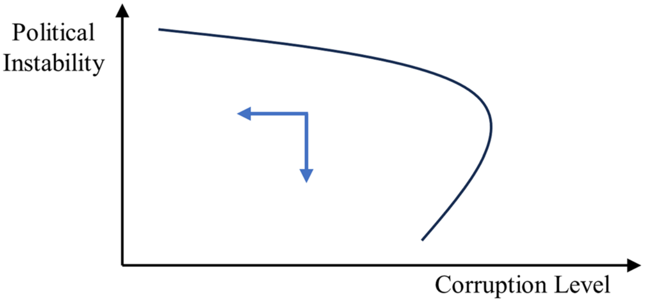

FIGURE 4 Welfare trade-offs. Indifference curve between corruption and political instability. Arrows indicate the direction under which welfare improves.

First, the speed at which resentment increases upon a corruption scandal, ω, appears as a good cultural trait: it lowers long-run corruption levels while it reduces political instability at the same time. Unfortunately, we cannot state whether it affects the set of initial states leading to exploding dynamics in one way or the other.

Instead, the impact of societal forgiveness or forgetfulness, ϕ, on welfare is multifaceted. On the one hand, it increases the long-term level of corruption activity. On the other hand, it decreases long-term resentment and thus political instability, but also directly increases the long-term hazard rate of political turnover. Additionally, ϕ narrows the spectrum of initial conditions leading to explosive dynamics.

More exogenous political instability, θ, is unambiguously bad for society's welfare if political instability is already high enough. It increases political instability directly and also through its effect on long-run resentment. It amplifies the range of initial states leading to exploding dynamics, yet it does not affect long-run corruption. However, when the average term duration is excessively extended, introducing factors that encourage political turnovers might be beneficial.

Finally, other parameters align with intuitive expectations. For example, increased rent opportunities from corruption, denoted by r, negatively affect long-term welfare. Conversely, higher penalties, denoted by s, and greater in-office benefits, denoted by m, are beneficial. Enhancing investigative and judicial resources, π, mitigates both corruption and instability, though its effect on the range of initial states leading to explosive dynamics is ambiguous. As mentioned above, the advantages of these parameters must, of course, be weighed against the implementation costs. Lastly, a more impatient briber, $\tilde{\rho}$, exacerbates political instability and expands the range of initial states leading to explosive dynamics.

7 | EXTENSIONS

Our model is subject to various interpretations and may respond to several modifications. In the ensuing discussion, we focus on two modifications that align most closely with the available empirical evidence: nonlinear growth rates for resentment and rent amount as a control variable.

Needless to say, the model admits more discussion than what we provide below. For instance, the briber could have a stronger antirejection technology other than "no more bribes in the future": she could credibly threaten to increase political instability, thereby increasing the chance of political turnover. Or, in a related fashion, the model may not take for granted that the briber can commit to an indefinite interruption of bribes. Additionally, the model could contain asymmetries between political parties, so that one of them holds an advantage, a valence, that makes it more likely to remain in power. The analysis of each variation would lead to a separate, entire piece of research.

7.1 | Nonlinear growth rate of resentment

One can also assume that the state of resentment influences its growth rate. Specifically, when resentment levels are sufficiently high, the rate of growth may decelerate, meaning new corruption scandals induce less additional outrage, whereas the converse could occur when resentment is low. This observation aligns with empirical evidence indicating that countries with prevalent corruption exhibit a higher tolerance towards corrupt practices compared to countries where corruption is less common (Barr & Serra, 2010; Cameron et al., 2009; Klašnja & Tucker, 2013). Further empirical support for this "corruption fatigue" hypothesis has been provided in recent studies by Pavão (2018) and Chang (2020).

Hence, rather than using $\dot{\lambda} = \lambda[-\phi + (\sigma + \phi)q]$, we propose an alternative law of motion for resentment:

$$\dot{\lambda} = \lambda^{\delta}[-\phi + (\sigma + \phi)q], \tag{33}$$

where $\delta \in (0, 1)$ captures the discussed concavity to the model, although it is pertinent to note that the model is functional for any positive δ. Thus, as δ decreases, the function λ^{δ} becomes more concave.

The model with this modification can be solved in the same fashion as the baseline model. We define

$$k_{\delta}(\lambda) := \tilde{\rho}\gamma(\lambda) + \phi\lambda^{\delta}\gamma'(\lambda). \tag{34}$$

The dynamics of the model remain qualitatively the same, leading to either a rapid convergence to a finite value of the resentment state or to explosive dynamics. When convergence occurs to a certain value λ_{δ}^{*}, this value satisfies both $k_{\delta}(\lambda_{\delta}^{*}) = 0$ and $k'_{\delta}(\lambda_{\delta}^{*}) < 0$.

We assume here that the steady state in the baseline model (when $\delta = 1$), is high enough, particularly $\lambda^{*} > 1$. We aim to see if the model with $\delta < 1$ close to 1 delivers a higher or lower steady state, as compared to the baseline model. Note that k_{δ} would be decreasing in δ, thus the steady state λ_{δ}^{*} would be decreasing in δ. Thus, $\delta < 1$ would deliver $\lambda_{\delta}^{*} > \lambda^{*}$. A slower speed of reaction to corruption scandals leads then to higher long-run resentment.

However, also the threshold value separating higher initial states leading to explosive dynamics from lower initial states leading to fast convergence, which we now denote here with $\hat{\lambda}_{\delta}$, will be decreasing in δ. Indeed, λ_{δ}^{*} must satisfy

$$M(\hat{\lambda}_{\delta}) = \mathrm{e}^{-\tilde{\rho}\tilde{T}\left(\hat{\lambda}_{\delta},\lambda_{\delta}^{*},\phi,\delta\right)}\frac{\phi}{\tilde{\rho}(\sigma + \phi)}\gamma\left(\lambda_{\delta}^{*}\right) - \int_{0}^{\infty}\mathrm{e}^{-\tilde{\rho}t}\gamma(\tilde{\lambda}(t))\,dt = 0, \tag{35}$$

where $\tilde{\lambda}(t) := \left[\hat{\lambda}_\delta^{1-\delta} + (1-\delta)\sigma t\right]^{\frac{1}{1-\delta}}$ and $\tilde{T}(\hat{\lambda}_\delta, \lambda_\delta^*, \phi, \delta) := \frac{\hat{\lambda}_\delta^{1-\delta} - \lambda_\delta^{*1-\delta}}{(1-\delta)\phi}$. Similarly to Equation (30), the first term in Equation (35) indicates the present value of payoffs if the briber makes λ decrease to λ_δ^* and the second term the present value of payoffs resulting from being active starting from $\hat{\lambda}_\delta$.

With a lower δ, λ increases more slowly towards λ_δ^*, thereby diminishing the first component's value. Additionally, due to this slower progression, the value of each $\tilde{\lambda}(t)$ remains below that of the baseline model's $\hat{\lambda}\,e^{\sigma t}$ for $\delta < 1$, leading to an increased value of the integral. Consequently, waiting for λ to reach its steady state becomes less advantageous, and continuous corruption activity becomes more convenient. Therefore, $M(\hat{\lambda}_\delta)$ diminishes as δ decreases, implying that $\hat{\lambda}_\delta$ also reduces with a decrease in δ. Relative to the baseline model, incorporating $\delta < 1$ broadens the spectrum of initial states that trigger explosive dynamics.

In summary, adopting the assumption of diminishing marginal effects of corruption on resentment results in a higher steady-state resentment level and broadens the range of resentment levels that precipitate unbounded political instability beyond what is suggested by the baseline model. Essentially, scenarios of “corruption fatigue” lead to increased persistence of corrupt activities and, paradoxically, cause the pervasiveness of political disaffection to grow over the long term to a point where the political system becomes entirely unreliable.

7.2 | Rent as a control variable

The model could also take a different perspective about the control variable, r. Since the briber is the true principal in the mechanism, then he could be thought of as being able to fully manage the bribing schedule by choosing the optimal level of the rent to be extracted each time. This is in line with the empirical literature finding evidence of increased corruption with the size of the possible rent from bribing. For example, Brollo et al. (2013) estimate an increase in local corruption following an increase in central government transfers. Furthermore, they find evidence of a nonlinear impact of larger rents on resentment, since the electoral punishment of disclosed corruption episodes is found to be lower when federal transfers are larger.

To capture those features, instead of fixing the instant rent opportunity to a single value r and then deciding on the probability of being active, the model could be modified by assuming that the briber chooses directly the size r of the corrupt activity within a domain $[0, \bar{r}]$. In other words, within this framework, the briber is presented with a range of possible rents from corruption and chooses the level that maximizes the present value of payoffs. Therefore, we are removing the briber's decision on the probability of being active, q, as in the baseline model, since at the chosen optimal rent level the briber will always be active.

With this approach, it is consequential to assume that the increase in resentment levels should depend on the size of corruption, as follows:

$$\dot{\lambda} = \lambda\left[\pi\bar{\omega}\, r^\delta - (1-\pi)\phi\right], \tag{36}$$

where δ measures the convexity of the dynamics of λ with respect r.

The Hamiltonian would become

$$H_r = e^{-\tilde{\rho}t}\gamma(r, \lambda) - \beta\lambda\left[\pi\bar{\omega}\, r^\delta - (1-\pi)\phi\right]. \tag{37}$$

Assuming $\delta \leqslant 1$ makes our model to be in line with the findings in Brollo et al. (2013), with a concave effect of disclosed corruption on electoral punishment in the size of federal transfers. With this assumption, the Hamiltonian would be convex in r, therefore the optimal path involves extreme choices (either no activity or $r = \bar{r}$), similar to the baseline model.[18]

The model provides further insights with $\delta > 1$. For instance, Vuković (2020) presents evidence supporting the nonlinear impact of corruption on reelection prospects, yet the effect occurs in the opposite direction compared to when $\delta \leqslant 1$. Specifically, voters tend to penalize corrupt actions only when the corruption is excessively widespread and apparent. Additionally, Vuković (2020) observes that corruption's positive influence on reelection predominantly occurs in smaller and mid-sized municipalities, where oversight is minimal and accountability is low. In contrast, for larger cities, the impact of corruption on reelection chances is negative, thereby affirming the overall nonlinear effect.

Assuming convexity in the effect of r on λ yields an interior solution for r, provided that δ is sufficiently high to ensure that $r^* < \bar{r}$. More in detail, with $\delta > 1$, and similarly to the baseline, the model may deliver either none, one, or two steady states. However, with none or one steady state the model predicts exploding dynamics, while if there are two steady states, one of them will be saddle-path stable, while the other will imply cyclical dynamics. The analysis of the dynamics of the model encompassing this extension is provided in the appendix.

8 | CONCLUSION

The analysis presented in this paper delves into the intricate relationship between corruption, societal resentment, and the political landscape through a comprehensive model, wherein a briber co-opts political parties, and the electorate accumulates resentment against a system corrupted by exposed malfeasance. Public resentment emerges as a pivotal force in the dynamics of corruption; it can burden incumbents and deter future corruption. Yet, this same resentment might also culminate in pathological political instability, which undermines social welfare.

The study yields several key insights summarized as follows.

Influence of societal forgiveness and vigilance: The model underscores the dual role of societal forgiveness and vigilance in shaping the steady-state levels of corruption and public resentment. Heightened vigilance and sensitivity to corruption scandals can deter corruption but do not directly alter the long-term resentment level, which is a novel insight indicating that the immediate public response to corruption does not necessarily translate into long-term societal attitudes. Conversely, societal forgiveness can lower the barriers to corruption by reducing the consequences of corrupt actions, thereby potentially raising the long-run levels of corruption. Surprisingly, faster forgiveness also helps prevent instability by making explosive dynamics of resentment and corruption less likely. This creates a policy dilemma: should we accept some instability to prevent endemic corruption?

External instability breeds corruption: Instability arising from factors *unrelated* to corruption enables the corrupt system to sustain itself in the long run, even amid high levels of resentment. Moreover, more external instability amplifies the set of initial resentment states that ignite a dynamics of exploding, unstoppable corruption activity. Concurrently, efforts to

[18] With $\omega = \bar{\omega}\, \bar{r}^{\delta}$.

broadly enhance political stability may have an unintended reduction in anticorruption anger and a lower risk of never-ending, fully-fledged corruption.

Welfare trade-offs: The paper proposes a welfare function to evaluate policy decisions. This function suggests that while removing corrupted governments is vital for societal well-being, *some degree of political instability (to be able to evaluate and punish ineffective policies) is also desirable. The ideal policy delicately balances these competing concerns.*[19]

Cultural traits towards corruption: The introduction of nonlinear resentment growth rates offers a more realistic portrayal of public responses to corruption, reflecting empirical observations across various governance contexts. This significantly enriches the model by accounting for public sensitivities to corruption, which vary based on historical exposure. The analysis reveals that cultural traits exhibiting high tolerance towards corrupt practices not only amplify the persistence of these activities but also, paradoxically, increase the pervasiveness of political disaffection over the long term, to the point where the political system becomes entirely unreliable.[20]

Rent-seeking behavior: If bribers can directly control the size of the gains from corruption, the model suggests dynamics where either there is no corruption or corruption operates at its maximum potential. This depends on factors like how strongly voters react against corruption and whether those reactions are more intense when the scale of corruption is blatant.

Policy implications and strategic responses: The analysis of policy parameters adapting to resentment levels highlights the strategic employment of policy tools to mitigate corruption and its effects. In periods of significant public discontent with the political system due to the spread of corruption scandals, allocating additional resources to combat corruption can help mitigate the resulting wave of political instability. An emergency policy that substantially raises the cost of engaging in corruption by allocating exceptional resources to investigative and judiciary powers upon exceeding a specific level of resentment provides a proactive approach to managing corruption. This is particularly vital in scenarios where corrupt entities disregard public sentiment, and the risk of corruption becoming rampant looms large. However, the implementation of such policies incurs significant costs, underscoring the importance of targeted and timely policy actions. This emergency policy may assist policymakers in deterring explosive dynamics, if it renders corruption activities unprofitable. On the one hand, this policy will not affect the long-run level of resentment. On the other hand, a briber, anticipating the possibility of such a policy, may find it less profitable to embark on a path of continuous corruption activity. This would eventually increase the threshold for explosive dynamics. The mere presence of this trigger policy reduces the range of initial resentment levels that catalyze active corruption, emphasizing its strategic implications.

In conclusion, the model presented and the subsequent analysis shed light on the complex dynamics at play in the realm of political corruption and public resentment. By carefully examining the effects of various parameters on corruption and societal welfare, this paper contributes to a deeper understanding of how corruption can be effectively managed and mitigated through thoughtful policy measures and an understanding of societal dynamics. The insights gained from this study not only advance theoretical knowledge but also offer practical guidance for policymakers seeking to navigate the challenges of corruption and political instability.

[19]On balancing corruption with policies in the environmental regulation context, see Ko et al. (2023).

[20]Similarly, on a related issue, Radhakrishnan (2022) shows that lobbying may lead to exploding dynamics in the misallocation of public expenditure if a very specific set of conditions is not met.

Finally, for future research, a potential modification of the model, beyond those already discussed, could consider that resentment impacts not only the likelihood of political turnover but also the bipartisan system itself. It is plausible that electoral turnout might significantly decrease as public resentment increases, creating opportunities for third parties (Carreras, 2017; Lucardie, 2000). Hence, the similarity in attitudes towards corruption between ruling and opposition parties may partly explain the trend towards instability in contemporary democracies.

AUTHOR CONTRIBUTIONS

All authors contributed equally to each section of the manuscript.

ACKNOWLEDGMENTS

We are grateful to all participants in the XIX Annual Conference of the Italian Society of Law & Economics (SIDE) in Brescia, the LXIV Annual Conference of the Italian Economic Association (SIE) in L'Aquila, the VIII Annual Conference of World Interdisciplinary Network for Institutional Research (WINIR) in Catania, the III Meeting in Economic Design in Padua and the seminar series at the Universidad Publica de Navarra for helpful comments and suggestions. The usual disclaimers apply. We thank the financial support from the Italian Prin2022 program (project code 2022HLPMKN), funded by the European Union—Next Generation EU. Open access publishing facilitated by Universita degli Studi di Messina, as part of the Wiley - CRUI-CARE agreement.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

ORCID

Maurizio Lisciandra http://orcid.org/0000-0002-9020-7106
Antonio Miralles Asensio http://orcid.org/0000-0001-8146-2140

REFERENCES

Anderson, C. J., & Tverdova, Y. V. (2003). Corruption, political allegiances, and attitudes toward government in contemporary democracies. *American Journal of Political Science*, *47*(1), 91–109. https://doi.org/10.1111/1540-5907.00007

Barr, A., & Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, *94*(11–12), 862–869. https://doi.org/10.1016/j.jpubeco.2010.07.006

Ben-Yashar, R., Krausz, M., & Nitzan, S. (2023). Sovereign debt assistance and democratic decision-making. *Journal of Public Economic Theory*, *25*(5), 930–943. https://doi.org/10.1111/jpet.12656

Brollo, F., Nannicini, T., Perotti, R., & Tabellini, G. (2013). The political resource curse. *American Economic Review*, *103*(5), 1759–1796. https://doi.org/10.1257/aer.103.5.1759

Cameron, L., Chaudhuri, A., Erkal, N., & Gangadharan, L. (2009). Propensities to engage in and punish corrupt behavior: Experimental evidence from Australia, India, Indonesia and Singapore. *Journal of Public Economics*, *93*(7–8), 843–851. https://doi.org/10.1016/j.jpubeco.2009.03.004

Cantú, F., & García-Ponce, O. (2015). Partisan losers' effects: Perceptions of electoral integrity in Mexico. *Electoral Studies*, *39*, 1–14. https://doi.org/10.1016/j.electstud.2015.03.004

Carreras, M. (2017). Institutions, governmental performance and the rise of political newcomers. *European Journal of Political Research*, *56*(2), 364–380. https://doi.org/10.1111/1475-6765.12181

Chang, E. C. (2020). Corruption predictability and corruption voting in Asian democracies. *Public Choice*, *184*(3–4), 307–326. https://doi.org/10.1007/s11127-019-00760-x

Chong, A., De La, L., De La, O. A. L., Karlan, D., & Wantchekon, L. (2015). Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice, and party identification. *The Journal of Politics*, *77*(1), 55–71. https://doi.org/10.1086/678766

Cooper, J. (2021). Political corruption cycles in democracies and autocracies: Evidence from micro-data on extortion in West Africa. *Quarterly Journal of Political Science*, *16*, 285–323. https://doi.org/10.1561/100.00019175

Dawid, H., & Feichtinger, G. (1996). On the persistence of corruption. *Journal of Economics*, *64*(2), 177–193. https://doi.org/10.1007/BF01250114

De Vries, C. E., & Solaz, H. (2017). The electoral consequences of corruption. *Annual Review of Political Science*, *20*, 391–408. https://doi.org/10.1146/annurev-polisci-052715-111917

Della Porta, D., & Vannucci, A. (2012). *The hidden order of corruption: An institutional approach*. Ashgate Publishing Ltd.

Dufwenberg, M., & Spagnolo, G. (2015). Legalizing bribe giving. *Economic Inquiry*, *53*(2), 836–853. https://doi.org/10.1111/ecin.12162

Farzanegan, M. R., & Witthuhn, S. (2017). Corruption and political stability: Does the youth bulge matter? *European Journal of Political Economy*, *49*, 47–70. https://doi.org/10.1016/j.ejpoleco.2016.12.007

Feitchinger, G., & Wirl, F. (1994). On the stability and potential cyclicity of corruption in governments subject to popularity constraints. *Mathematical Social Sciences*, *28*, 113–131. https://doi.org/10.1016/0165-4896(94)00750-0

Ferraz, C., & Finan, F. (2008). Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, *123*(2), 703–745. https://doi.org/10.1162/qjec.2008.123.2.703

Figueroa, V. (2020). Political corruption cycles: High-frequency evidence from Argentina's doibooks scandal. *Comparative Political Studies*, *54*, 482–517.https://doi.org/10.1177/0010414020938102

Gaspar, A., Giommoni, T., Morelli, M., & Nicoló, A. (2021). *Corruption and extremism* [BAFFI CAREFIN Centre Research Paper, No. 2021–163].

Giannoccolo, P., & Lisciandra, M. (2019). Political corruption and capture of the minority. *Economics & Politics*, *31*(3), 485–510. https://doi.org/10.1111/ecpo.12137

Klašnja, M., & Tucker, J. A. (2013). The economy, corruption, and the vote: Evidence from experiments in Sweden and Moldova. *Electoral Studies*, *32*(3), 536–543. https://doi.org/10.1016/j.electstud.2013.05.007

Ko, C. Y., Shen, B., & Zhang, X. (2023). Can corruption encourage clean technology transfer? *Journal of Public Economic Theory*, *25*(3), 459–492. https://doi.org/10.1111/jpet.12627

Kostadinova, T. (2009). Abstain or rebel: Corruption perceptions and voting in East European elections. *Politics & Policy*, 37(4), 691–714. https://doi.org/10.1111/j.1747-1346.2009.00194.x

Lee, J.-H., Iwasa, Y., Dieckmann, U., & Sigmund, K. (2019). Social evolution leads to persistent corruption. *Proceedings of the National Academy of Sciences*, *116*(27), 13276–13281. https://doi.org/10.1073/pnas.1900078116

Lisciandra, M., Miralles Asensio, A., & Monteforte, F. (2024). Search and matching in political corruption. *The BE Journal of Economic Analysis & Policy*, *24*(1), 69–96. https://doi.org/10.1515/bejeap-2023-0022

Lucardie, P. (2000). Prophets, purifiers and prolocutors: Towards a theory on the emergence of new parties. *Party Politics*, *6*(2), 175–185. https://doi.org/10.1177/1354068800006002003

Morris, S. D., & Klesner, J. L. (2010). Corruption and trust: Theoretical considerations and evidence from Mexico. *Comparative Political Studies*, *43*(10), 1258–1285. https://doi.org/10.1177/0010414010369072

Pavão, N. (2018). Corruption as the only option: The limits to electoral accountability. *The Journal of Politics*, *80*(3), 996–1010. https://doi.org/10.1086/697954

Pellegrini, L., & Gerlagh, R. (2008). Causes of corruption: A survey of cross-country analyses and extended results. *Economics of Governance*, *9*(3), 245–263. https://doi.org/10.1007/s10101-007-0033-4

Peters, J. G., & Welch, S. (1980). The effects of charges of corruption on voting behavior in congressional elections. *American Political Science Review*, *74*(3), 697–708. https://doi.org/10.2307/1958151

Radhakrishnan, R. (2022). Public expenditure allocation, lobbying, and growth. *Journal of Public Economic Theory*, *24*(4), 756–780. https://doi.org/10.1111/jpet.12570

Rinaldi, S., Feichtinger, G., & Wirl, F. (1998). Corruption dynamics in democratic societies. *Complexity*, *3*, 3–64. https://doi.org/10.1002/(SICI)1099-0526(199805/06)3:5<53::AID-CPLX10>3.0.CO;2-L

Vuković, V. (2020). Corruption and re-election: How much can politicians steal before getting punished? *Journal of Comparative Economics*, *48*(1), 124–143. https://doi.org/10.1016/j.jce.2019.09.002

Zhang, Y., & Kim, M.-H. (2018). Do public corruption convictions influence citizens' trust in government? The answer might not be a simple yes or no. *The American Review of Public Administration*, *48*(7), 685–698. https://doi.org/10.1177/0275074017728792

How to cite this article: Lisciandra, M., Miralles Asensio, A., & Monteforte, F. (2024). Corruption dynamics and political instability. *Journal of Public Economic Theory*, *26*, e12712. https://doi.org/10.1111/jpet.12712

APPENDIX A: PROOF OF LEMMA 1

Let $\alpha \coloneqq \beta\lambda$. Applying Pontryagin's optimality condition for λ and its law of motion yields

$$\dot{\alpha} = \dot{\beta}\lambda + \beta\dot{\lambda} = q\mathrm{e}^{-\tilde{\rho}t}\lambda\gamma'(\lambda). \tag{A1}$$

Differentiating the condition $g = 0$ gives

$$-\tilde{\rho}\mathrm{e}^{-\tilde{\rho}t}\gamma(\lambda) + \mathrm{e}^{-\tilde{\rho}t}\gamma'(\lambda)\lambda(-\phi + (\sigma + \phi)q) - \dot{\alpha}(\sigma + \phi) = 0. \tag{A2}$$

Substituting $\dot{\alpha}$ and rearranging leads to the condition $k = 0$. Given that k may have zero, one, or two roots and considering k is continuous in λ, it follows that λ assumes a unique value λ^* over the interval (T, T'). □

APPENDIX B: PROOF OF REMARK 1

(Monotonicity)

Suppose the contrary—that there exists a state value λ' and distinct moments t and t' such that, under an optimal bribing schedule, $\lambda(t) = \lambda' = \lambda(t')$, with $\lambda(\cdot)$ increasing at t and decreasing at t'. This implies $q(t) > q(t')$. Given that the history of previous actions and states does not influence the set of available actions at any future moment (which spans the interval $[0, 1]$), the optimal bribing schedule from any moment forward relies solely on the current state variable λ. Consequently, the bribing schedule post-t' would remain optimal if initiated from t, and vice versa. Therefore, for a given resentment level λ', the briber faces indifference between two distinct q values at t. This indifference extends to an open neighborhood (T, T') around t, due to the continuity of λ with respect to time. However, as delineated in Lemma 3, such conditions necessitate that λ remains constant over the interval, leading to a contradiction. □

APPENDIX C: PROOF OF LEMMA 3

Let us examine the behavior of λ^* with respect to a parameter u, starting from the condition

$$\frac{\partial k}{\partial \lambda^*}\frac{\partial \lambda^*}{\partial u} + \frac{\partial k}{\partial u} = 0. \tag{C1}$$

This relationship can be reformulated as

$$\frac{\partial \lambda^*}{\partial u} = -\frac{\frac{\partial k}{\partial u}}{\frac{\partial k}{\partial \lambda^*}}. \tag{C2}$$

Given $\frac{\partial k}{\partial \lambda^*} < 0$, it follows that the sign of $\frac{\partial \lambda^*}{\partial u}$ aligns with the sign of $\frac{\partial k}{\partial u}$. In a stable equilibrium, where $\gamma(\lambda^*) > 0$ and $\gamma'(\lambda^*) < 0$, it can be deduced from Equation (24) that λ^* increases with $\tilde{\rho}$ and decreases with ϕ, while it is invariant to changes in ω.

Further analysis requires examining the specific functional form of $k(\lambda^*)$, given by

$$k(\lambda^*) = \tilde{\rho}A + \frac{B}{d}x(cz^*) = 0, \tag{C3}$$

where $x(cz^*) = [-\tilde{\rho} + (\tilde{\rho} - \phi)cz + \phi(cz)^2]$ is negative.[21] This configuration suggests that r and s directly influence A—positively for r and negatively for s, leading to straightforward implications for λ^* in response to variations in r and s.

Concerning the derivative $\frac{\partial k}{\partial \pi}$, a decrease in A and an increase in $\frac{B}{d}$ occur, which, when combined with a negative multiplier $x(cz^*)$, results in the derivative being negative. Regarding $\frac{\partial k}{\partial \theta}$, a decrease in $\frac{B}{d}$, once again multiplied by a negative value, indicates a positive sign for the derivative. Conversely, an increase in $\frac{B}{d}$ due to changes in m—multiplied by a negative number—yields a negative sign for $\frac{\partial k}{\partial m}$. Lastly, $\frac{\partial k}{\partial \rho}$ combines a positive term A with a negative term derived from the positive derivative of $\frac{B}{d}$ with respect to ρ and the negative expression $x(cz^*)$, rendering the sign of this derivative ambiguous. □

APPENDIX D: PROOF OF LEMMA 4

We define $T\left(\hat{\lambda}, \lambda_1^*, \phi\right)$ as the duration required to reduce the resentment level from $\hat{\lambda}$ to λ_1^*. Assuming $\hat{\lambda}$ is well defined, we analyze the function $M(\hat{\lambda})$ as follows:

$$\begin{aligned} M(\hat{\lambda}) &= \mathrm{e}^{-\tilde{\rho}T\left(\hat{\lambda},\lambda_1^*,\phi\right)}\frac{\phi}{\tilde{\rho}(\sigma+\phi)}\gamma\left(\lambda_1^*\right) - \int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma(\hat{\lambda}\,\mathrm{e}^{\sigma t})\,dt \\ &= \left(\frac{\lambda_1^*}{\hat{\lambda}}\right)^{\tilde{\rho}/\phi}\frac{\phi}{\tilde{\rho}(\sigma+\phi)}\gamma\left(\lambda_1^*\right) - \int_0^\infty \mathrm{e}^{-\tilde{\rho}t}\gamma(\hat{\lambda}\,\mathrm{e}^{\sigma t})\,dt = 0. \end{aligned} \tag{D1}$$

The derivative $M'(\hat{\lambda}) < 0$ indicates a crucial point for comparative statics analysis (see footnote).

Given that $T\left(\hat{\lambda}, \lambda_1^*, \phi\right)$ decreases with ϕ and λ_1^* also decreases with ϕ, it follows that M increases with ϕ. Consequently, we deduce that $\hat{\lambda}$ increases with ϕ.

[21] Here, recall that $\frac{B}{d} = \frac{\rho m}{(1-\pi)(\rho+2\theta)}$.

An increase in π or ω has an ambiguous effect on M, and consequently, on the level of $\hat{\lambda}$. On the one hand, it reduces $q^* = \frac{\phi}{\sigma+\phi}$. On the other hand, it accelerates the resentment state in the integral, reducing the integral's value and inducing a positive effect on M.

The impact of an increase in $\tilde{\rho}$ on $\hat{\lambda}$ is negative. Notably, the derivative of $(\lambda_1^*)^{\tilde{\rho}/\phi}\gamma(\lambda_1^*)$ with respect to λ_1^* is zero, indicating that ρ has no indirect effect on M through its influence on λ_1^*. The partial derivative of M with respect to $\tilde{\rho}$ is given by

$$\begin{aligned}\frac{\partial M}{\partial \tilde{\rho}} &= -\left(\tilde{\rho} + \frac{1}{\tilde{\rho}}\right)\left(e^{-\tilde{\rho}T\left(\hat{\lambda},\lambda_1^*,\phi\right)}\frac{\phi}{\tilde{\rho}(\sigma+\phi)}\gamma\left(\lambda_1^*\right)\right) \\ &\quad +\tilde{\rho}\int_0^{\infty} e^{-\tilde{\rho}t}\gamma(\hat{\lambda}\,e^{\sigma t})\,dt \\ &= -\tilde{\rho}M(\hat{\lambda}) - e^{-\tilde{\rho}T\left(\hat{\lambda},\lambda_1^*,\phi\right)}\frac{\phi}{\tilde{\rho}^2(\sigma+\phi)}\gamma\left(\lambda_1^*\right) < 0.\end{aligned} \tag{D2}$$

Marginal variations of other variables influence $\hat{\lambda}$ only through their effect on the evaluation of γ along the integral. This is because the derivative of $(\lambda_1^*)^{\tilde{\rho}/\phi}\gamma(\lambda_1^*)$ with respect to λ_1^* is zero.

Recalling the effects of the parameters on γ discussed at the end of Section 3, we find that $\hat{\lambda}$ decreases with an increase in r, increases with an increase in s and m, and decreases with an increase in θ. The reaction of $\hat{\lambda}$ to a variation in ρ remains unclear. □

APPENDIX E: DYNAMICS WITH *r* AS A CONTROL VARIABLE

For $\delta > 1$, an interior path of r as a continuous function of $\beta\lambda$ would be characterized by

$$e^{-\tilde{\rho}t}\left[(1-\pi) - \pi s - \frac{\pi}{1-\pi}s\right] - \beta\lambda\pi\bar{\omega}\,\delta r^{\delta-1} = 0. \tag{E1}$$

Differentiating this expression and using the first order condition for the state variable, which yields $(\dot{\beta\lambda}) = e^{-\tilde{\rho}t}\lambda\frac{\partial\gamma}{\partial\lambda}$, we find

$$[-\tilde{\rho}r + (1-\delta)\dot{r}]r^{-\delta}\left[(1-\pi) - \pi s - \frac{\pi}{1-\pi}s\right] = \lambda\frac{\partial\gamma}{\partial\lambda}\pi\bar{\omega}\,\delta. \tag{E2}$$

A stationary state can be found by making both $\dot{\lambda}$ and $\dot{r}$ equal to zero. Imposing $\dot{\lambda} = 0$ yields a phase line parallel to the horizontal axis at the level of long-run corruption activity equal to

$$r^* = \left(\frac{(1-\pi)\phi}{\pi\bar{\omega}}\right)^{1/\delta}, \tag{E3}$$

which is intuitively increasing in ϕ and decreasing in all other arguments. Since $\frac{\partial\dot{\lambda}}{\partial r} = \delta\lambda\bar{\omega}\,r^{\delta-1} > 0$ horizontal arrows take the directions depicted in Figure E1.

Using (E2), the $\dot{r} = 0$ phase line reads

$$\dot{r} = \frac{r}{1-\delta}\left[\tilde{\rho} + r^{\delta-1}\lambda\gamma'(\lambda)\frac{\pi\bar{\omega}\,\delta}{A_r}\right] = 0, \tag{E4}$$

where $A_r \equiv \left[(1-\pi) - \pi s - \frac{\pi}{1-\pi}s\right]$. Therefore, for nontrivial values of r, the second phase line will depend on the function:

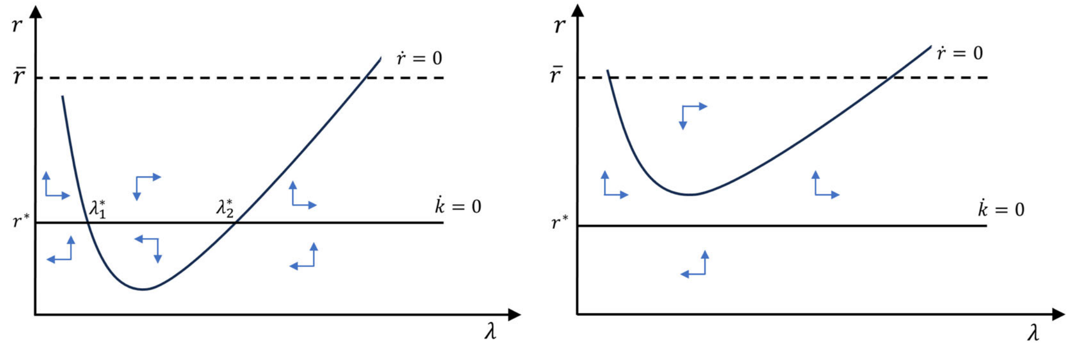

FIGURE E1 Dynamics with rent as a control variable. Phase diagrams between resentment and rent from bribing as a control variable. Right panel, case with two equilibria; left panel, case with no equilibria.

$$D(r,\lambda) = \left[\tilde{\rho} + r^{\delta-1}\lambda\gamma'(\lambda)\frac{\pi\bar{\omega}\,\delta}{A_r}\right] = 0, \tag{E5}$$

which implicitly defines a function of the form $r = r'(\lambda)$ which is either increasing or decreasing in the phase diagram depending on the sign of the following:

$$r'(\lambda) = -\frac{D_\lambda}{D_r} \tag{E6}$$

with

$$D_r = (\delta - 1)\lambda\gamma'(\lambda)\frac{\pi\bar{\omega}\,\delta}{A_r}r^{\delta-2} < 0, \tag{E7}$$

$$D_\lambda = r^{\delta-1}\frac{\pi\bar{\omega}\,\delta}{A_r}[\gamma'(\lambda) + \lambda\gamma''(\lambda)] > 0, \tag{E8}$$

where the sign in the second derivative depends on the sign of the term in squared brackets:

$$[\gamma'(\lambda) + \lambda\gamma''(\lambda)] = -\frac{Bc(c - d\lambda)}{(c + d\lambda)^3}, \tag{E9}$$

which is positive iff $\lambda > \frac{c}{d} = \frac{\rho+\theta}{\pi}$. If that is the case, and remembering that $\frac{\partial\dot{r}}{\partial\lambda} = D_\lambda > 0$, it implies that the $\dot{r} = 0$ locus is increasing for all $\lambda > \frac{c}{d}$ and decreasing for all $\lambda < \frac{c}{d}$, thus taking a U-shaped form. Furthermore, it has an asymptote in the vertical axis since $r \to \infty$ when $\lim_{\lambda\to 0}$.

However, the level of resentment in the steady state will depend on whether the two phase lines cross each other or not. This will depend on the parameters. The two main cases are depicted in Figure E1. In the left panel of Figure E1 we plot the case in which the two phase lines cross. If so, they will cross in two steady states $\lambda_1^* < \lambda_2^*$. Differently from the baseline model, in this framework, the direction of the arrows indicates that λ_2^* will be saddle-path stable, while λ_1^* admits cyclical dynamics.

On the other hand, if the $\dot{r} = 0$ locus always lies above the $\dot{\lambda} = 0$ phase line, as depicted in the right panel of Figure E1, then we will always have exploding dynamics. The same can be said in the particular case in which the two-phase lines are tangent in one point: the corresponding unique steady state will be unstable.