

# Electoral Transparency and Sequential Voter Rationality: A Dynamic Game-Theoretic Analysis of Implementing a Political Rating Agency

Akhenaton IZU-MAKONGO  
Université Grenoble Alpes

**Abstract:** Conventional electoral theory posits that transparency enhances accountability by enabling voters to make informed, rational decisions. However, empirical evidence from diverse national contexts reveals a persistent disconnect: the disclosure of corruption frequently fails to precipitate significant electoral repercussions. This apparent paradox has led some scholars to posit voter irrationality. This paper challenges this interpretation, contending instead that the absence of electoral sanctions against corrupt politicians arises from the inherent complexities of strategic environments, characterized by information asymmetry and ambiguous signaling. We develop a formal model employing dynamic game theory to analyze the strategic interactions between politicians and voters who update their beliefs through Bayesian learning. Our model demonstrates that electoral accountability emerges incrementally, contingent upon sequences of credible signals across multiple electoral cycles. To mitigate the pervasive challenges of adverse selection and moral hazard, we propose the establishment of an independent Political Rating Agency (PRA). This agency would furnish credible, standardized, and readily interpretable information. The efficacy of such a PRA is critically contingent upon its perceived credibility and the electorate's propensity to integrate its ratings into their decision-making calculus.

**Keywords:** Electoral Accountability, Principal-Agent Problem, Bayesian Learning, Electoral Transparency, Political Rating Agency

**JEL Code:** D72, D83, H11

## 1. INTRODUCTION

The intricate nexus between transparency, accountability, and voter behavior presents a formidable challenge to traditional electoral theory. While seminal works posit that informed electorates hold elected officials accountable (Fearon, 1999; Ferejohn, 1986; Besley, 2006), and conversely, that political ignorance impedes effective electoral functioning (Brennan, 2016; Somin, 2016), empirical evidence offers a more variegated perspective. Although certain studies document instances where voters demonstrably punish corruption (Krause & Mendez, 2009; Klasna, 2015; Wood & Grose, 2022), a growing body of research indicates a weakened correlation between information dissemination and electoral accountability. Field experiments conducted across diverse geopolitical landscapes—including Brazil (Ferraz & Finan, 2008; De Figueiredo et al., 2023), India (Banerjee et al., 2011), Mexico (Chong et al., 2015), Kenya (Palmer & McClendon, 2025), the Netherlands (Rienks, 2023), and the USA (Dulay & Lee, 2024; Breitenstein & Hernández, 2024)—consistently reveal that the exposure of corrupt practices does not reliably translate into significant electoral consequences. As Pande (2011) astutely observed, the mere provision of information is seldom sufficient to engender electoral accountability.

This empirical divergence has fueled discourse concerning voter "irrationality," with scholars such as Achen and Bartels (2016) suggesting that electoral choices are frequently driven by socio-political identity rather than reasoned deliberation. However, research by Lee et al. (2017) on Taiwanese elections offers a counterpoint, positing that while voters can exhibit rationality, the propensity for irrational choices escalates in higher-stakes electoral contests. This evolving comprehension of voter cognition transcends the paradigm of "rational ignorance" (Downs, 1957), moving towards the concept of "rational irrationality" (Caplan, 2001; Paulson, 2023). This framework posits that the maintenance of factually erroneous beliefs can be instrumentally rational, particularly when the personal cost of such

beliefs is negligible and they confer psychological or emotional benefits. This underscores the complex interplay of informational inputs, psychological predispositions, and political decision-making processes.

Transparency, in isolation, proves insufficient for achieving electoral accountability. This insufficiency can be attributed to three principal mechanisms. Firstly, voters frequently employ cognitive heuristics-biases and shortcuts—to navigate the inherent complexities of political information. This can precipitate motivated reasoning, wherein individuals selectively seek information that corroborates pre-existing beliefs while dismissing contradictory evidence, even concerning favored candidates (Bullock, 2011; Cotter et al., 2020; Nasr, 2021). Furthermore, voters are susceptible to the fundamental attribution error, exhibiting a propensity to misattribute negative outcomes to politicians' actions rather than to exogenous factors, such as economic downturns (Ashworth et al., 2018; Healy & Malhotra, 2013; Weitz-Shapiro & Winters, 2017). Glaeser and Ponzetto (2017) argue that this attributional bias has profound implications for institutional design, potentially leading to an insufficient demand for a free press, an excessive demand for authoritarian governance, and a preference for electing ostensibly "honest" leaders over implementing systemic reforms to combat corruption. Moreover, a significant field experiment in Mexico by Chong et al. (2015) indicated that the revelation of local government corruption paradoxically decreased voter turnout, suggesting that such information engendered cynicism or disillusionment rather than targeted electoral punishment. This finding resonates with other research highlighting voters' difficulties in processing complex, contradictory, or technical information (Arceneaux & Johnson, 2013; Metag and Gurr, 2022). Neuroscientific research on political cognition corroborates that the processing of political information activates emotional and identity-based neural circuits, thereby impeding purely rational belief updating (Amodio & Jost, 2012).

Secondly, social and partisan identities often supersede ethical considerations. Voters typically prioritize their party's success over abstract notions of good governance. Party affiliation functions as a potent filter through which accusations of corruption are evaluated, often leading to the dismissal of such claims when directed at one's own political faction (Campbell et al., 1960; Anduiza, Gallego, & Munoz, 2013; Iyengar et al., 2019; Kumar Jha, 2023). This tendency is further reinforced by research demonstrating that voters frequently respond more robustly to clientelism and appeals to ethnic identity than to policy platforms (Wantchekon, 2003). This prioritization of identity over policy is a well-documented phenomenon within the study of social identity and party loyalty (Levitsky & Ziblatt, 2018). This dynamic exacerbates political polarization, entrenching voters within their respective partisan and social enclaves. Crucially, the mere dissemination of additional information does not inherently bridge this divide. Studies have indicated that increased information consumption can, in fact, intensify polarization, particularly among highly partisan individuals (Hetherington, 2009; Kalmoe, 2020; Hannon, 2022). This polarization also manifests in a "winner-loser" gap, wherein citizens whose parties are in power report significantly higher levels of satisfaction with democratic processes than their counterparts in opposition parties (Martini & Quaranta, 2019; Nadeau et al., 2023; Eck & Paulis, 2025).

Thirdly, political elites actively engage in information manipulation rather than the transparent provision of factual data. Incumbents often leverage state resources for clientelistic purposes, exert control over media outlets to shape public narratives (Djankov et al., 2003; Besley & Prat, 2006; Van Remoortere and Vliegenthart, 2025), and deploy sophisticated disinformation campaigns that inundate the information ecosystem with "noisy signals." This obfuscates truth and cultivates voter skepticism (Guriev & Treisman, 2020; Bartoloni et al, 2022). The proliferation of social media has amplified this challenge, with studies such as Allcott and Gentzkow (2017) demonstrating the potent influence of "fake news" in shaping political beliefs and exacerbating media polarization. The pervasive use of automation (e.g., bot networks) and the dissemination of conspiracy theories or rumors online further contribute to this informational degradation (Gentzkow et al., 2016, Ferrara et al, 2020).

This article challenges the prevailing notion of voter irrationality. Rather than signifying a deficit in reasoning, the observed lack of electoral sanctions for suboptimal performance may represent a logical consequence of operating within a complex strategic environment. We present a formal model designed to explicate this dynamic, featuring:

- **Voters as Bayesian Learners:** Voters possess prior beliefs that are updated through the processing of noisy signals, received with a certain probability. They iteratively refine their understanding of political performance based on available, albeit imperfect, information.
- **Politicians as Strategic Actors:** Motivated by career advancement, politicians operate within a repeated game framework characterized by moral hazard. Their prospects for re-election are contingent upon past performance, with first-term politicians exhibiting heightened effort due to their greater electoral vulnerability.

Our model simulates the intricate interplay between these two principal actors across a complete electoral cycle—encompassing pre-election, election, governance, and re-election phases—utilizing a dynamic, sequential game with imperfect information. We demonstrate how voters employ Bayesian learning to refine their beliefs over time, effectively navigating the inherent uncertainty of political signaling.

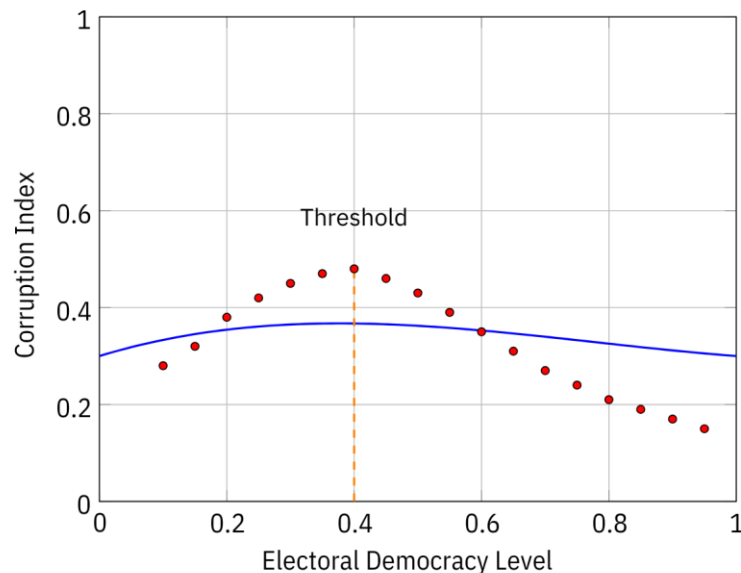
The introduction of a Political Rating Agency fundamentally alters the information environment for both actors. By providing credible, independent pre-election assessments of candidates, grounded in their educational background, experience, and ethical track record, the agency actively combats misinformation and empowers voters to make more informed decisions. This independent vetting process serves to mitigate "adverse selection"—the risk of electing unsuitable candidates due to a paucity of reliable information. Beyond initial candidate evaluation, the agency continues to monitor the performance and decisions of elected officials, regularly updating their ratings based on rigorous investigations. This sustained scrutiny addresses "moral hazard"—the risk of politicians behaving irresponsibly once in office. The agency's consistent oversight fosters accountability and promotes responsible governance. Furthermore, the agency also tackles the "attribution problem" by discerning whether positive or negative economic outcomes are attributable to a politician's specific actions or are merely the result of chance. This nuanced analytical capacity enables voters to more equitably assess performance, thereby diminishing the inherent advantage often enjoyed by incumbents. Crucially, the agency's effectiveness is inextricably linked to its perceived credibility, independence, and the public's trust. Maintaining these foundational attributes is paramount for the agency to achieve its objectives and exert a positive influence on the political landscape.

## 2. METHODS

### 2.1. Theoretical Framework: Sequential Rationality and Cumulative Knowledge

McMann et al. (2017) investigated the relationship between corruption and electoral democracy using data from 173 countries spanning 1900 to 2012. Their analysis identified a nonlinear pattern: in countries with nascent democratic institutions, an initial increase in corruption is observed during the transition towards greater electoral democracy. However, beyond a specific threshold, further democratization typically correlates with a more consistent reduction in corruption.

Figure 1. The relationship between corruption and electoral democracy



This empirical observation suggests the role of voter learning over time, suggesting that when a country introduces or reintroduces electoral mechanisms (regardless of their initial quality), corruption may initially escalate. However, as the quality of electoral processes improves, corruption tends to decline.

The interaction between voters and political elites (both incumbents and candidates) is conceptualized as a sequential game under conditions of imperfect information. In this framework, each player strategically adjusts their actions and anticipates the resultant payoffs over time. This scenario closely mirrors the principal-agent problem prevalent in corporate governance, where shareholders (principals) delegate operational management to executives (agents) who oversee daily activities, ideally aligned with objectives set by the board. In this corporate context, principals engage in continuous monitoring of executives through independent audits and the analysis of various performance indicators. Given the sequential nature of the game and the presence of imperfect information, principals may not ascertain the alignment of executive interests with their own through early audits alone. Over time, as more indicators become available, principals can form more accurate assessments. This establishes a continuous learning process wherein principals update their information regarding performance and subsequently inform their strategic decisions.

Voters, similarly, engage in a Bayesian learning process (Seeberg et al., 2016; Hill, 2017; Vaeth, 2024). They are not epistemically inert; rather, they commence with prior beliefs about a candidate or governmental actions, shaped by factors such as party affiliation, historical performance, or reputation (Alesina & Rosenthal, 1995; Tappin et al., 2020). Each new piece of information functions as a signal, which voters integrate with their prior beliefs to form updated beliefs. In contexts of information asymmetry, these signals are inherently noisy, being imperfectly observed through mechanisms such as audits or media reports (Vaeth, 2024). For instance, an economic slowdown (a negative signal) could stem from corruption (indicating a "bad" elite type) or from an exogenous global recession. Conversely, the

construction of a new school (a positive signal) might be the product of competent governance or the result of inflated construction contracts involving illicit payments. Moreover, unlike corporate shareholders who often rely on expert teams to analyze decisions with clearly defined indicators, the average voter faces significant challenges in fully interpreting each incoming signal (Nai, 2019). Their belief updates are consequently slower and more iterative, demanding greater cognitive effort. Empirical evidence from Stoetzer et al. (2024) highlights "cautious Bayesian learning," wherein voters update imperfectly their beliefs by incorporating new information. Hill (2017) conducted an experiment with 990 participants from Amazon's Mechanical Turk (MTurk) and found that subjects learn as cautious Bayesians, updating their beliefs at approximately 73 percent of the rate of perfect Bayesian application. He also observed that for information consistent with prior beliefs, subject learning was statistically indistinguishable from perfect Bayesian updating. However, for inconsistent information, learning was significantly less than perfect. Despite these biases, beliefs did not exhibit polarization.

This interaction is not a singular event but rather a sequential game (North, Wallis & Weingast; 2009). A voter's electoral decision is predicated on the beliefs they have updated at that specific juncture (sequential rationality). Guarino (2021) illustrates sequential rationality in noncooperative strategic interactions through an iterative method termed "conditional B-Dominance," which posits that "a rational player selects from her strategy set the one that maximizes her subjective expected utility, i.e., maximizing utility according to beliefs formed about opponents' strategies." Cumulative knowledge (Izzo et al., 2018) is constructed across multiple electoral cycles. A single pre-election audit may yield an ambiguous signal, whereas a historical record of audits and political actions spanning a decade offers a more discernible view of a politician's underlying type, facilitating more accurate belief updates and, consequently, more informed electoral sanctions. If voters do not act upon corruption disclosures, this does not necessarily imply they are "polluting the polls" (Brennan, 2009), but rather that they are imperfectly updating their beliefs based on the noisy information provided.

Empirical studies tracking this learning process over time lend support to these findings. Ferraz and Finan (2011), in a crucial follow-up to their 2008 study, found that voters do indeed learn, exhibiting higher rates of electoral punishment for mayors in areas where corruption detection is robust, facilitated by local media and judicial systems. Wood & Grose (2022) utilized a natural experiment that randomly assigned varying levels of campaign transparency to legislators, discovering that incumbents facing greater transparency were more likely to retire. This enhanced transparency aids voters in discerning traits such as honesty or corruption. A field experiment conducted in 354 subnational constituencies in Uganda by Grossman, Michelitch & Prato (2023) provides strong evidence that sustained transparency enhances electoral accountability, even in weakly institutionalized settings. They conclude that information regarding politicians' performance must be disseminated early, regularly, and predictably throughout their tenures.

## 2.2. The Base Model

This model offers a powerful framework for understanding the complex dynamics of political decision-making. It integrates principles of dynamic game theory, signaling theory, and political agency to analyze the strategic interactions between political elites and the population they govern, particularly under conditions of information asymmetry, moral hazard, and adverse selection. By incorporating mixed strategies, the model realistically reflects the often unpredictable and ambiguous nature of elite behavior in the real world. Crucially, the model demonstrates how the perceived future value of holding office, represented by the discount factor ( $\delta$ ), significantly influences elite choices, driving them toward either public-spirited investment or self-enriching corruption. This framework provides valuable insights into the incentives that shape political behavior and the resulting consequences for public welfare.

The game features two main players: the population (principal) and the political elite (agent). The population (P) aims to improve public welfare, whereas the elites (E) focus on maximizing their personal benefits from office and private

gain. Nature determines an elite's type  $\theta \in \{\theta_H, \theta_L\}$  where  $\theta_H$  is a high type who is honest and supportive of development, with prior probability  $p_H = P(\theta_H)$ ;  $\theta_L$  is a low-type who is corrupt and self-serving, with prior probability  $P(\theta_L) = 1 - p_H$ . This unobservable "type" is a classic feature of adverse selection problems.

If elected, an elite chooses an action  $a \in \{I, E, M_I, M_E\}$ . The action I corresponds to pure investment, which incurs a cost  $C_I$  to the elite, yielding a benefit  $G > 0$  to the population. In contrast, action E means pure embezzlement, providing a benefit  $B_E > 0$  to the elite but no benefit to the population. The elite can also adopt mixed strategies:  $M_I$  represents an investment-dominant mixed strategy, investing with probability  $q_I > 0.5$  and embezzling with probability  $1 - q_I$ ; conversely,  $M_E$  is an embezzlement-dominant mixed strategy, embezzling with probability  $q_E > 0.5$  and investing with probability  $1 - q_E$ . Meanwhile, the population adopts electoral strategies: for the election decision,  $e \in \{1, 0\}$  refers to either electing or not electing the incumbent; likewise, for the re-election decision,  $r \in \{1, 0\}$  refers to re-electing or not. The game unfolds over several stages representing the electoral cycle.

Before the election (stage 0), nature selects the elite's type  $\theta$ . The elite decide on the level of electoral fraud  $f$  and make campaign promises  $\rho$ . The population observes noisy signals  $s_0$  regarding the elite's type  $\theta$ . During the first election (stage 1), the population votes, and the elite wins with a probability  $p_w(f, \mu_1, \rho)$ . Following this, in the action phase (stage 2), if elected, the elite selects an action  $a$  from a set of actions. The population observes a noisy signal  $s_2$  about the chosen action  $a$ . Finally, during re-election (stage 3), the population updates its belief to  $\mu_2$ , based on  $s_2$  and decides whether to re-elect the elite.

### 2.2.1. Notation and assumptions

- $\mu_0 = P(\theta = \theta_H)$ : The population's prior belief that the elite is of high type before any interaction occurs.
- $s_0$ : A noisy signal about the elite's type  $\theta$  observed by the population before the first election. The quality of the signal is determined by the likelihood  $P(s_0/\theta)$ .
- $f \in (0, 1)$ : The level of electoral fraud chosen by the elite, where 0 means no fraud and 1 means maximum fraud.
- $\rho$ : campaign promises made by the elite, which can be a vector or a scalar quantity.
- $p_w(f, \mu_1, \rho)$ : The probability that the incumbent elite wins the first election. This probability increases with the level of fraud  $f$ , the population's updated belief  $\mu_1$  about the elite being a high type, and the attractiveness of promises  $\rho$ :
- $a \in \{I, E, M_I, M_E\}$ : elite action — pure investment in public goods, pure embezzlement of public funds, mixed strategy-investment dominant (the elite invests primarily in public goods but engages in some level of corruption with probability  $\alpha$ .) and mixed strategy-embezzlement dominant (the elite mainly embezzles funds but invests a fraction in public goods with probability  $\beta$  )
- $s_2$ : A noisy signal about the elite's chosen action  $a$ , with likelihood  $P(s_2/a, \theta)$ . This signal helps the population to imperfectly learn about the elite's behavior in office.
- $\delta \in (0, 1)$ : The discount factor, indicating how much elites and the population value future payoffs compared with current payoffs. A higher  $\delta$  means that future outcomes (such as re-election) are more relevant.
- $U_E^t, U_P^t$ : payoffs to the elite and population, respectively at stage  $t$

### 2.2.2. Stages of the Game

At stage 0, before the election, nature draws the elite's type with  $P(\theta = \theta_H) = \mu_0$  reflecting the initial level of uncertainty. The elite strategically selects the level of electoral fraud  $f$  and makes campaign promises  $\rho$ . The population observes a noisy signal  $s_0$ , and the conditional probability density  $P(s_0/\theta) = \phi_\theta(s_0)$  depends on the elite's type, implying that some signals are more likely for certain types. The elite's payoff at this stage is given by:

$$U_E^0(f, \rho, \theta) = B_0(\rho, \theta) + b_f * f - C(f)$$

Here  $B_0(\rho, \theta)$  is the benefit from campaign promises, based on the elite's type,  $b_f * f$  captures the benefit derived from electoral fraud, with marginal benefit  $b_f$ , and  $C(f)$  is a convex cost function that accounts for logistical, reputational, or penalty costs, such that  $C'(f) > 0$  and  $C''(f) > 0$ .

Correspondingly, the population's payoff at this stage relies on their expectations of the elite's future behavior:

$$U_P^0(s_0, \rho) = \mathbb{E}_{\theta/s_0}[V_p(\theta, \rho)]$$

Here  $V_p(\theta, \rho)$  is the expected utility from the elite's promises, and  $\mathbb{E}_{\theta/s_0}$  denotes the expected value based on the posterior belief about the elite's type  $\theta$ .

The population updates its belief via Bayes' rule as follows:

$$\mu_1(\theta_H/s_0) = \frac{P(s_0/\theta_H)\mu_0}{P(s_0/\theta_H)\mu_0 + P(s_0/\theta_L)(1 - \mu_0)}$$

This updated belief  $\mu_1$  reflects the probability that the elite is of type  $\theta_H$  given the signal  $s_0$ .

At stage 1, during the first election, the population votes based on their updated belief  $\mu_1$  and campaign promises  $\rho$ . The elite's probability of winning is as follows:

$$p_w(f, \mu_1, \rho) = p_0(\mu_1, \rho) + \beta f \cdot f$$

where  $p_0(\mu_1, \rho)$  is a baseline probability depending on population belief  $\mu_1$  (a higher belief in a good type improves chances) and the attractiveness of promises, and where  $\beta f \cdot f$  captures the additional winning probability due to electoral fraud, with  $\beta f > 0$ .

The elite decides on  $f$  and  $\rho$  to maximize their expected payoff:

$$V_E^1 = \max_{f, \rho} \{U_E^0(f, \rho, \theta) + \delta \cdot p_w(f, \mu_1, \rho) \cdot V_E^2(\theta)\}$$

where  $V_E^2(\theta)$  is the elite's continuation value (expected future utility) from holding office.

Simultaneously, the population's expected payoff is:

$$V_P^1 = p_w(f, \mu_1, \rho) \cdot \mathbb{E}_{\theta \sim \mu_1}[U_P^2(\theta)] + (1 - p_w(f, \mu_1, \rho)) \cdot U_P(\text{Challenger})$$

where  $U_P^2(\theta)$  is the expected utility from the elite's future action, and where  $U_P(\text{Challenger})$  is the expected utility with a probability  $1 - p_w$  if the challenger wins.

At stage 2, the elite chooses an action  $a \in \{I, E, M_I, M_E\}$ . The population observes a noisy signal  $s_2$  related to the elite's action with likelihood  $P(s_2/a, \theta) = \psi_{a, \theta}(s_2)$ .

The elite payoff in this stage depends on the expected return of the chosen strategy:

$$U_E^2(a, \theta) = \mathbb{E}[r_a \theta] + \delta \cdot V_E^3(\mu_2)$$

Here  $r_a(\theta)$  depends on the chosen strategy:

$$r_a(\theta) = \begin{cases} r_I(\theta) & \text{if } a = I \\ q_I \cdot r_I(\theta) + (1 - q_I) \cdot r_E(\theta) & \text{if } a = M_I \\ q_E \cdot r_E(\theta) + (1 - q_E) \cdot r_I(\theta) & \text{if } a = M_E \\ r_E(\theta) & \text{if } a = E \end{cases}$$

$V_E^3(\mu_2)$  is the continuation value based on the belief updated after observing  $s_2$ . Payoffs from mixed strategies are weighted averages of payoffs from pure strategies, either investment or embezzlement.

Based on the observed signal and the (inferred or announced) action  $a$ , the population further updates its belief via Bayes' rule as:

$$\mu_2(\theta_H/s_2, a) = \frac{P(s_2/a, \theta_H)\mu_1(\theta_H)}{P(s_2/a, \theta_H)\mu_1(\theta_H) + P(s_2/a, \theta_L)(1 - \mu_1(\theta_H))}$$

The belief  $\mu_1(\theta_H)$  (posterior from stage 0) now serves as the prior for this stage.

The population's payoff is:

$$U_P^2(a, \theta) = v_a - d_a(\theta)$$

where  $v_a$  and  $d_a(\theta)$  depend on the chosen strategy:

$$v_a = \begin{cases} v_I & \text{if } a = I \\ q_I \cdot v_I - (1 - q_I) \cdot d_E(\theta) & \text{if } a = M_I \\ (1 - q_E) \cdot v_I - q_E \cdot d_E(\theta) & \text{if } a = M_E \\ -d_E(\theta) & \text{if } a = E \end{cases}$$

Where  $v_a$  is the direct value of action  $a$  to the population, and  $d_a(\theta)$  represents losses from corruption or embezzlement.

At stage 3, the population decides whether to re-elect the incumbent based on the updated belief  $\mu_2$ . The decision rule is binary:

$$R = \begin{cases} 1, & \text{if } \mu_2 \geq \bar{\mu} \\ 0 & \text{Otherwise} \end{cases}$$

where  $\bar{\mu}$  is a threshold belief for re-election, possibly indicating the expected quality of a generic challenger.

The elite's payoff in this stage is as follows:

$$V_E^3(\mu_2) = R \cdot B_3(\theta)$$

with  $B_3(\theta)$  as the benefit from holding office again

The population's payoff is:

$$V_P^3 = R \cdot U_P(re - elected, \theta) + (1 - R) \cdot U_P(challenger)$$

If the incumbent is reelected, the population receives utility  $U_P(re - elected, \theta)$ , which depends on the incumbent's type. If not, a challenger is elected, yielding  $U_P(challenger)$ .

Regarding equilibria, a pooling equilibrium occurs when both types choose the same strategy  $a$ , leading to  $\mu_1 \approx \mu_2$  where beliefs do not update meaningfully. If they pool on actions, this results in no information about the elite's type for the population. This happens when signals  $s_0$  and  $s_2$  are noisy (high variance), fraud costs  $C_f$  are low, and corruption benefits are high. Conversely, a separating equilibrium occurs when different types of elites select distinct

strategies.  $\theta_H$  might choose pure investment  $I$ , whereas  $\theta_L$  opts for pure embezzlement or a more embezzlement-focused mixed strategy. As strategies differ, the population can accurately infer the elite's type after observing the action or update their beliefs more significantly even with noisy signals ( $|\mu_1 - \mu_2|$  large). This situation occurs when signals are clear, fraud costs are high, and elite values for future payoffs are strong ( $\delta$  large).

### 2.3. Introduction of A Political Rating Agency (PRA)

Informed decision-making is the bedrock of a healthy democracy, yet voters today often struggle to access clear and trustworthy information. The current media and NGO landscape frequently falls short, leaving citizens adrift in a sea of complex reports and conflicting narratives. To address this critical gap, we propose the establishment of independent Political Rating Agencies (PRAs).

Modeled after established financial rating institutions, PRAs would provide objective assessments of political actors – including parties, candidates, and elected officials. These ratings would focus on essential areas such as integrity, competence, policy performance, and governance. Instead of requiring voters to sift through dense reports and biased news, PRAs would distill complex political information into easily understandable, standardized ratings. This crucial process would effectively separate genuine performance from external influences. For example, a mayor's handling of a pandemic could be assessed independently of the pandemic's overall impact, clearly distinguishing the effects of policy choices from the influence of unpredictable events. This approach empowers voters with the clarity needed to select effective leaders.

Political Rating Agencies would play a vital role in informing voters and strengthening democratic processes through several key functions:

- **Providing Accessible and Credible Information:** PRAs would meticulously collect and analyze comprehensive data on political candidates, encompassing their backgrounds, voting records, policy positions, campaign promises, and any history of corruption. This in-depth research would offer voters readily accessible, fact-based insights that are otherwise difficult and time-consuming to gather independently, enabling more informed decisions based on verifiable data.
- **Reducing Adverse Selection:** By publicly rating candidates before elections, PRAs would help voters distinguish between qualified individuals and those who may be problematic. This transparency would minimize the risk of electing unsuitable candidates – those who might be corrupt, incompetent, or misaligned with the electorate's values. This pre-election evaluation acts as a crucial safeguard against "bad actors" and promotes accountability.
- **Mitigating Moral Hazard:** Through ongoing monitoring and rating of incumbents during their terms, PRAs would serve as continuous checks. This incentivizes politicians to remain diligent, act transparently, and deliver public goods, as poor ratings could jeopardize their re-election prospects.
- **Enhancing Voter Decision-Making:** PRAs empower voters by simplifying complex political information. Easy-to-understand scorecards or grades would distill governance data into accessible formats, enabling all voters, regardless of their political expertise, to make more informed choices at the ballot box.
- **Supporting Institutional Trust:** By operating as non-partisan, expert bodies, PRAs would foster greater institutional trust. In an era of polarized media and partisan division, they would offer a credible, objective source of information, thereby strengthening the foundations of democratic governance.

By issuing and regularly updating ratings in real-time, PRAs would create valuable long-term performance records. This would allow voters to observe not just snapshots of a candidate's current standing, but also their performance trends over time, accelerating knowledge accumulation. This approach directly addresses challenges identified by Chong et al. (2015), where voters often dismiss last-minute information as partisan attacks or react with cynicism.

Furthermore, simple rating systems offer powerful, easily digestible guides that significantly reduce the cost of acquiring political information. Voters gain access to condensed, actionable insights, empowering them to make more informed decisions. This directly challenges the "rational irrationality" hypothesis proposed by Paulson (2023), which suggests voters avoid the time-consuming and inconvenient process of gathering political information. With readily available ratings, voters would no longer need to expend significant effort to stay informed, fostering a more engaged and empowered electorate.

## 2.4. Augmented model

The model expands to include a political rating agency (PRA) that provides **public, credible ratings**  $R_t$  on elites and political parties at various stages. These ratings serve as additional signals about elite types  $\theta$  and their past actions. The PRA's ratings have a precision parameter  $\sigma_R^{-2}$ , improving the overall clarity of the signals available to the public.

We integrate the PRA's rating into our baseline model, introduced in section II.2. Below is the stage-by-stage analysis.

At stage 0, the population observes both a private noisy signal  $s_0$  with probabilities  $P(s_0/\theta) = \phi_\theta(s_0)$  and a public rating  $R_0$  from the PRA, generated as:

$$R_0 = \tilde{r}_0(\rho, \theta) + \varepsilon_0,$$

where  $\varepsilon_0 \sim \mathcal{N}(0, \sigma_R^2)$  is normally distributed noise with a mean of 0 and variance of  $\sigma_R^2$ , representing imprecision in the PRA rating. A smaller  $\sigma_R^2$  indicates a more precise rating. The rating  $R_0$  relies on the elite's promises  $\rho$  and their true type  $\theta$  (for example, a high-type might make more credible promises, resulting in a better  $r_0$ ). The population now combines private signal  $s_0$  and the public PRA's rating  $R_0$  to update their beliefs. A favorable rating will increase  $\mu_1$ , especially if the rating is precise (low  $\sigma_R^2$ ):

$$\mu_1(\theta_H/s_0, R_0) = \frac{\phi_H(s_0) \cdot \psi_H(R_0) \cdot \mu_0}{\sum_{\theta' \in \{\theta_H, \theta_L\}} \phi_{\theta'}(s_0) \cdot \psi_{\theta'}(R_0) \cdot P(\theta')}$$

where  $\psi_{\theta'}(R_0)$  is the PRA rating probability under type  $\theta$

The elite's payoff in this stage is based on the benefits of campaigning, the level of fraud, and the costs tied to fraud:

$$U_E^0(f, \rho, \theta) = B_0(\rho, \theta) + b_f f - C(f) - \gamma \cdot \mathbb{I}[R_0 < R^*]$$

where  $\gamma$  represents a parameter for the reputational cost or penalty linked to a poor PRA rating. This introduces a new incentive for elites to make credible promises or behave in ways that garner positive PRA ratings.

The expected utility for the public at this stage is now based on an expectation concerning the more informed posterior belief  $\mu_1(\theta_H/s_0, R_0)$ :

$$U_P^0(s_0, R_0, \rho) = \mathbb{E}_\theta[V_p(\theta, \rho)]$$

At stage 1, the election outcome probability also depends on an election-related PRA rating  $R_1$ , defined as :

$$p_w = p_0(\mu_1, \rho, R_1) + Bf$$

where  $R_1$  is the PRA rating on campaign/election conduct including fraud and promises:

$$R_1 = \tilde{r}_1(f, \rho, \theta) + \varepsilon_1, \quad \varepsilon_1 \sim \mathcal{N}(0, \sigma_R^2)$$

A positive  $R_1$  (for example, for clean campaigns and realistic promises) can directly increase the baseline probability of winning  $p_0$ . This suggests that elites might rely less on fraud  $f$  if they can secure a good  $R_1$ .

The population updates its beliefs by incorporating  $R_1$  as follows:

$$\mu_2(\theta_H/R_1) = \frac{\psi_H(R_1) \cdot \mu_1(\theta_H)}{\psi_H(R_1) \cdot \mu_1(\theta_H) + \psi_L(R_1) \cdot (1 - \mu_1(\theta_H))}$$

At this stage, the elite maximizes the expected payoff:

$$V_E^1 = \max_{f, \rho} \{U_E^0(f, \rho, \theta) + \delta \cdot p_w(f, \mu_1, \rho) \cdot V_E^2(\theta) - \gamma \cdot \mathbb{I}[R_1 < R^*]\}$$

Elites now also take into account the reputational cost  $\gamma \cdot \mathbb{I}[R_1 < R^*]$  linked to a poor rating regarding their campaign conduct. This diminishes the incentive for high fraud rates or extravagant promises if they risk being flagged by the PRA.

The population's expected payoff similarly hinges on belief  $\mu_2$ , refined by the PRA's ratings as follows:

$$V_P^1 = p_w \cdot \mathbb{E}_{\theta \sim \mu_2} [U_P^2(\theta)] + (1 - p_w) \cdot U_P(\text{Challenger})$$

At stage 2, the elite's action choices and the public's observations are enhanced by a public rating  $R_2$

$$R_2 = \tilde{r}_2(a, \theta) + \varepsilon_2, \quad \varepsilon_2 \sim \mathcal{N}(0, \sigma_R^2)$$

The population beliefs update accordingly as follows:

$$\mu_3(\theta_H/s_2, R_2, a) = \frac{\phi_H(s_2) \cdot \psi_H(R_2) \cdot \mu_2(\theta_H)}{\sum_{\theta' \in \{\theta_H, \theta_L\}} \phi_{\theta'}(s_2) \cdot \psi_{\theta'}(R_2) \cdot \mu_2(\theta')}$$

The elite's payoff also includes reputational penalties:

$$U_E^2(a, \theta) = r_a(\theta) + \delta \cdot V_E^3(\mu_2) - \gamma \cdot \mathbb{I}[R_2 < R^*]$$

The population's payoff remains:

$$U_P^2(a, \theta) = v_a - d_a(\theta)$$

At stage 3, the reelection decision depends on  $\mu_3(\theta_H/s_2, R_2, a)$  exceeding  $\bar{\mu}$ , and the PRA releases a cumulative rating:

$$R_3 = \tilde{r}_3(\text{history}) + \varepsilon_3, \quad \varepsilon_3 \sim \mathcal{N}(0, \sigma_R^2)$$

The elite's payoff is:

$$V_E^3(\mu_2) = \mathbb{I}[\mu_3 \geq \bar{\mu}] \cdot B_3(\theta) - \gamma \cdot \mathbb{I}[R_3 < R^*]$$

And the population's payoff is:

$$V_P^3 = \mathbb{I}[\mu_3 \geq \bar{\mu}] \cdot U_P^{\text{re-elected}}(\theta) + \mathbb{I}[\mu_3 < \bar{\mu}] \cdot U_P^{\text{challenger}}$$

### 3. RESULTS

#### 3.1. Impact on noisy signals and beliefs

The core impact of the political rating agency (PRA) is to reduce noise in the information system by combining the signals and ratings at each stage, leading to improved signal precision:

$$\sigma_{combined,t}^{-2} = \sigma_{s_t}^{-2} + \sigma_{R_t}^{-2}$$

This noise reduction has significant effects:

- **Sharper Bayesian Updates:** Beliefs  $(\mu_1, \mu_3)$  become more responsive to actual elite behavior and type. The population can more accurately distinguish  $\theta_H$  from  $\theta_L$ .
- **Mitigation of Adverse selection:** With clearer information before elections (from  $R_0, R_1$ ), the population is better prepared to select high-type elites.
- **Mitigation of Moral Hazard:** Knowing that their actions  $a$  will be assessed by the PRA (via  $R_2, R_3$ ) and that these evaluations are relatively transparent, elites have less motivation to engage in hidden opportunistic actions.

Our model suggests that a credible PRA can shift political competition from a contest of vague promises and manipulation to a more policy-based contest, as elites are incentivized to earn a high rating through tangible performance.

#### 3.2. Reputational costs and shifts in elite strategies

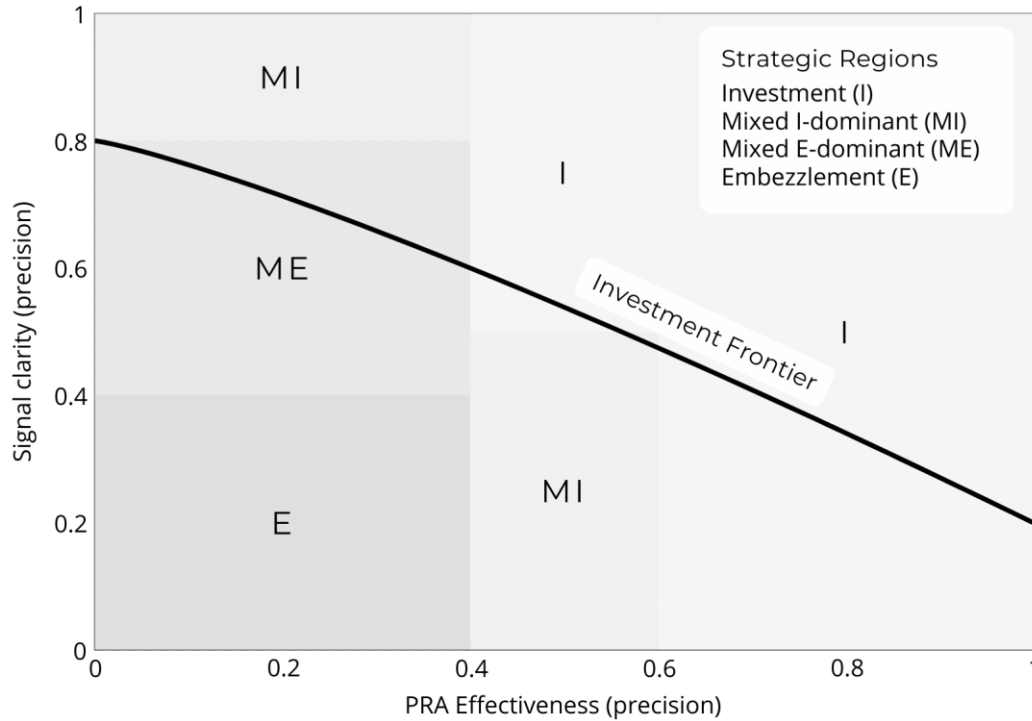
The introduction of clear penalties  $\gamma$  for poor PRA ratings  $\gamma \cdot \mathbb{I}[R_t < R^*]$  directly creates **reputational costs** for elites. A minimum empirical rating,  $R^*$ , must be achieved within a given period to qualify for election. Once you obtain less than this threshold, these costs can result in:

- Lowered public esteem.
- Reduced chances of re-election or future political success.
- Increased scrutiny and criticism from media and civil society.

Understanding these reputational costs encourages the elite to modify their strategies:

- **Before election:** Make more credible promises and possibly reduce electoral fraud to avoid negative  $R_0$  and  $R_1$  ratings.
- **Action phase:** Shift from embezzlement-focused  $(E, M_E)$  strategies toward investment-focused  $(I, M_I)$  strategies to prevent poor  $R_2$  ratings and improve the chances of a good cumulative  $R_3$ .

Figure 2: Elite Strategy Decision Map



This decision map shows how PRA effectiveness and signal precision interact to determine elite strategies. When the signals are unclear and the PRA ineffective, embezzlement remains the dominant strategy. As signals become clearer and the PRA becomes more effective, elites are more likely to select investment-dominant strategies, either pure Investment or mixed-dominant investment strategies.

### 3.3. Impact on equilibria

This makes **separating equilibria** more likely, where high-type elites genuinely invest and receive recognition for it (via PRA ratings and later through population beliefs), whereas low-type elites are either deterred from harmful actions or are more easily identified and penalized. Conditions for pooling equilibria (where bad types can hide) become less likely when a precise and influential PRA operates.

## 4. DISCUSSION

### 4.1. Institutional Credibility and Feasibility

For a political rating agency to be effective, it must maintain strong credibility and independence. Its institutional design should incorporate a transparent methodology, diversified funding sources to avoid capture by special interests, and governance structures that ensure independence (Besley & Burgess, 2002). The comparison with credit rating agencies, despite their known flaws (Sinclair, 2005), shows that reputation mechanisms can uphold institutional value over time. Research on the effects of fact-checkers and nonpartisan information sources shows that voters can respond positively to such intermediaries (Nyhan & Reifler, 2015; Graves, 2016).

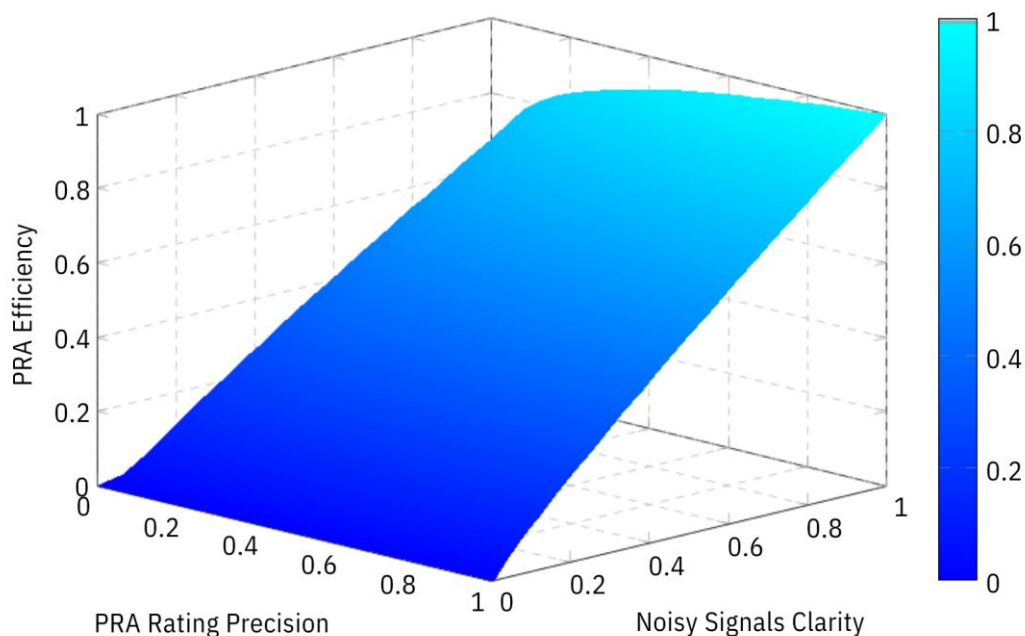
### 4.2. Effect of population reactivity

The effectiveness of a PRA does not depend only on its existence and precision, it also depends on how much the public trusts and reacts to its ratings. If  $\omega \in (0, 1)$  represents how much weight the public gives to PRA ratings, the effective likelihood for belief updates changes to:

$$P^{eff}(R_t/\theta) = P(R_t/\theta)^\omega \cdot P(s_t/\theta)^{1-\omega}$$

- **Unreactive population** ( $\omega \rightarrow 0$ ): The population largely ignores or distrusts the PRA's rating, relying mostly on noisy private signals  $s_t$ . In this case, the PRA's impact is minimal, and issues of adverse selection and moral hazard continue to a great extent.

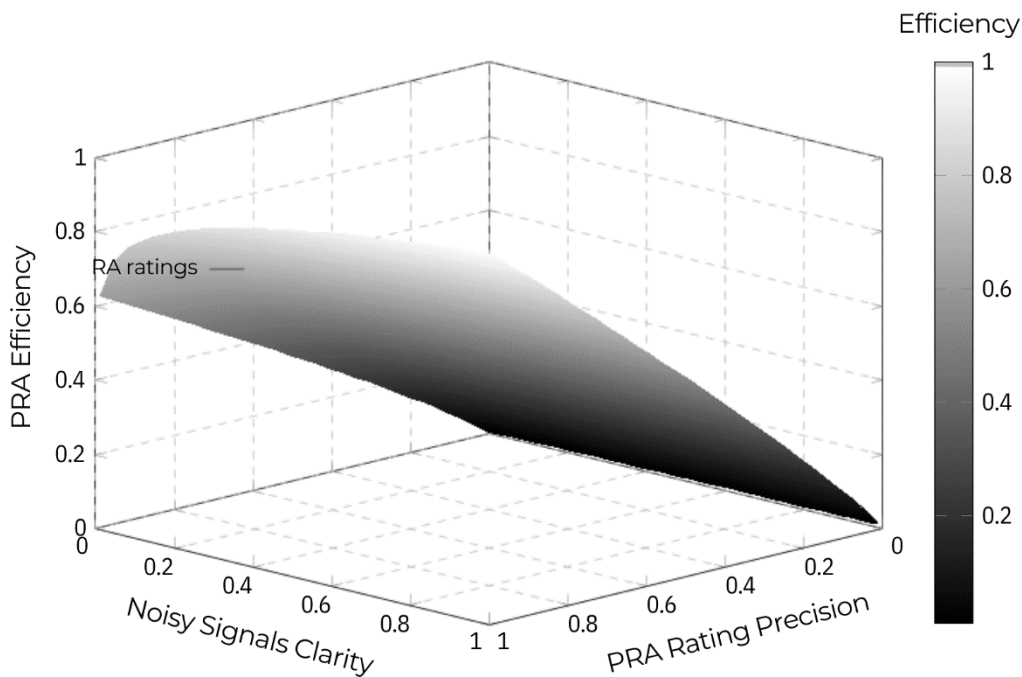
Figure 3: PRA Efficiency with minimal public trust



The above graph illustrates that the PRA efficiency changes most dramatically with changes in noisy signals along with the y-axis, and less so with changes in PRA rating along the x-axis. Factors contributing to low reactivity might include perceived bias in the PRA, a lack of understanding of its ratings, or strong partisan loyalties that overshadow objective assessments.

- **Reactive population** ( $\omega \rightarrow 1$ ): The population trusts and heavily relies on the PRA's ratings. Private noisy signals  $s_t$  are mostly disregarded. In this scenario, the influence of the political rating agency is at its peak.

Figure 4: PRA Efficiency with maximal public trust



This figure demonstrates that when public reactiveness is high, PRA efficiency is more sensitive to the its rating and less sensitive to the noisy signals. The elite faces strong pressure to earn good ratings, promoting a shift towards investment-focused strategies and accountability.

## 5. CONCLUSION

The consistent failure of voters to punish corrupt incumbents is not necessarily a sign of widespread irrationality. Instead, it likely represents a rational, albeit imperfect, outcome within a complex strategic environment characterized by manipulated and unreliable information. Our model frames voters as Bayesian learners engaged in a continuous, sequential interaction with strategic politicians. Within this framework, accountability isn't an immediate reaction to new information, but rather a gradual learning process reliant on a consistent flow of credible, high-quality signals over multiple election cycles.

The introduction of a political rating agency (PRA) offers a promising institutional solution to the information deficits that undermine electoral accountability. Acting as a professional signal processor, a PRA can mitigate both adverse selection before elections and moral hazard during governance. By providing voters with clear, standardized, and longitudinal data, the agency facilitates effective Bayesian updating, reduces the cognitive burden on voters, and counters elite manipulation. However, the PRA's success hinges on two critical factors: maintaining its credibility and independence, and cultivating voter trust and reliance on its ratings. These conditions are essential for the agency to achieve its intended impact.

Further research should empirically test the model's predictions and explore the practical challenges of implementing such an agency across diverse political landscapes. This paper ultimately argues that strengthening electoral accountability relies less on criticizing voter irrationality and more on establishing robust information institutions that empower voters' inherent capacity for Bayesian reasoning. By shifting the focus from voter deficits

to institutional design, this paper provides a blueprint for building the informational foundations that make democratic accountability possible.

## V. REFERENCES

Achen, C.H. & Bartels, L.M. (2016). *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton University Press.

Alesina, A. & Rosenthal, H. (1995). *Partisan Politics, Divided Government, and the Economy*. Cambridge University Press.

Allcott, H. & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), pp.211-236. <https://doi.org/10.1257/jep.31.2.211>.

Amodio, D.M. & Jost, J.T. (2012). Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion*, 36, pp.55-64. <https://doi.org/10.1007/s11031-011-9260-7>.

Anduiza, E., Gallego, A. & Muñoz, J. (2013). Turning a blind eye: Experimental evidence of partisan bias in attitudes toward corruption. *Comparative Political Studies*, 46(12), pp.1664-1692. <https://doi.org/10.1177/0010414013495769>.

Arceneaux, K. & Johnson, M. (2013). *Changing Minds or Changing Channels? Partisan News in an Age of Choice*. University of Chicago Press.

Ashworth, S., Bueno de Mesquita, E. & Freidenberg, A. (2018). Learning about voter rationality. *American Journal of Political Science*, 62(1), pp.37-54. <https://doi.org/10.1111/ajps.12324>.

Banerjee, A., Kumar, S. and Pande R. (2011). Do Informed voters make better choices? Experimental evidence from Urban India (Unpublished manuscript), Massachusetts Institute of Technology; Carnegie Mellon University; Harvard University.

Bartoloni, D., Sacchi, A., Scalera, D. & Zazzar, A. (2022). Voters' distance, information bias and politicians' salary. *Italian Economic Journal*. <https://doi.org/10.1007/s40797-022-00183-6>.

Besley, T. (2006). *Principled Agents? The Political Economy of Good Government*. Oxford University Press.

Besley, T. & Burgess, R. (2002). The political economy of government responsiveness: Theory and evidence from India. *Quarterly Journal of Economics*, 117(4), pp.1415-1451. <https://doi.org/10.1162/003355302320935061>.

Besley, T. & Prat, A. (2006). Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review*, 96(3), pp.720-736. <https://doi.org/10.1257/aer.96.3.720>.

Brennan, J. (2009). Polluting the polls: When citizens should not vote. *Australasian Journal of Philosophy*, 87(4), pp.535-549. <https://doi.org/10.1080/00048400903174007>.

Brennan, J. (2016). *Against Democracy*. Princeton University Press.

- Breitenstein, S. & Hernández, E. (2024). Too crooked to be good? Trade-offs in the electoral punishment of malfeasance and corruption. *European Political Science Review*, 17(1), pp.61-79. <https://doi.org/10.1017/S175577392400016X>.
- Bullock, J.G. (2011). Elite influence on public opinion in an informed electorate. *American Political Science Review*, 105(3), pp.496-515. <https://doi.org/10.1017/S0003055411000204>.
- Campbell, A., Converse, P.E., Miller, W.E. & Stokes, D.E. (1960). *The American Voter*. University of Chicago Press.
- Caplan, B. (2001). Rational irrationality and the microfoundations of political failure. *Public Choice*, 107(3), pp.311-331.
- Chong, A., et al. (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), pp.55-71. <https://doi.org/10.1086/677468>.
- Cotter, R.G., Lodge, M. & Vidigal, R. (2020). The boundary conditions of motivated reasoning. In: R. K. Tolbert (ed.) *The Oxford Handbook of Electoral Persuasion*, pp.66-87.
- De Figueiredo, M.F.P., Hidalgo, D.F. & Kasahara, F.Y. (2023). When do voters punish corrupt politicians? Experimental evidence from a field and survey experiment. *British Journal of Political Science*, 53(2), pp.728-739. <https://doi.org/10.1017/S0007123421000347>.
- Djankov, S., McLiesh, C., Nenova, T. & Shleifer, A. (2003). Who owns the media? *The Journal of Law and Economics*, 46(2), pp.341-382. <https://doi.org/10.1086/375341>.
- Downs, A. (1957). *An Economic Theory of Democracy*. Harper Collins.
- Dulay, D.C. & Lee, S. (2024). Sorry not sorry: Presentational strategies and the electoral punishment of corruption. *Electoral Studies*, 92, p.102683. <https://doi.org/10.1016/j.electstud.2023.102683>.
- Dunning, T., et al. (2019). Voter information campaigns and political accountability: Cumulative findings from a preregistered meta-analysis of coordinated trials. *Science Advances*, 5(7), eaaw2612. <https://doi.org/10.1126/sciadv.aaw2612>.
- Eck, B. & Paulis, E. (2025). Defending the status quo or seeking change? Electoral outcome, affective polarization, and support for referendums. *British Journal of Political Science*, 55, May 2025.
- Ejdemyr, S., Kramon, E. & Robinson, A.L. (2018). Segregation, ethnic favoritism, and the strategic targeting of local public goods. *Comparative Political Studies*, 51(9), pp.1111-1143. <https://doi.org/10.1177/0010414018758751>.
- Fearon, J.D. (1999). Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance. In: A. Przeworski, S.C. Stokes & B. Manin, eds. *Democracy, Accountability, and Representation*. Cambridge University Press, pp. 55-97.
- Ferejohn, J. (1986). Incumbent Performance and Electoral Control. *Public Choice*, 50 (1-3), 5-25.

- 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), pp.703-745. <https://doi.org/10.1162/qjec.2008.123.2.703>.
- Ferraz, C. & Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4), pp.1274-1311. <https://doi.org/10.1257/aer.101.4.1274>.
- Ferrara, E., Chang, H., Chen, E., Muric, G. & Patel, J. (2020). Characterizing social media manipulation in the 2020 US presidential election. *First Monday*, 25(11). <https://doi.org/10.5210/fm.v25i11.11431>.
- Gentzkow, M., Shapiro, J.M. & Stone, D.F. (2016). Media bias in the marketplace: Theory. In: N.L. Rose, ed. *Handbook of Media Economics*, vol. 1, pp.623-645. North-Holland.
- Glaeser, E.L. & Ponzetto, G.A.M. (2017). Fundamental errors in the voting booth. NBER Working Paper Series, No. 23683.
- Grossman, G., Michelitch, K. & Prato, C. (2023). The effect of sustained transparency on electoral accountability. *American Journal of Political Science*, 67(1), pp.20-30. <https://doi.org/10.1111/ajps.12787>.
- Graves, L. (2016). *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*. Columbia University Press.
- Guarino, P. (2021). Sequential rationality and ordinal preferences. Unpublished manuscript. SSRN. Available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3621347](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3621347).
- Guriev, S. & Treisman, D. (2020). A theory of international autocracy. *Journal of Public Economics*, 190, 104158. <https://doi.org/10.1016/j.jpubeco.2020.104158>.
- Hannon, M. (2022). Are knowledgeable voters better voters? *Politics, Philosophy, Economics*. <https://doi.org/10.1177/1470594X211065080>.
- Healy, A. & Malhotra, N. (2013). Retrospective voting reconsidered. *Annual Review of Political Science*, 16, pp.285-306. <https://doi.org/10.1146/annurev-polisci-100711-135654>.
- Hetherington, M.J. (2009). Putting polarization in perspective. *British Journal of Political Science*, 39(2), pp.413-448.
- Hill, S.J. (2017). Learning together slowly: Bayesian learning about political facts. *Journal of Politics*, 79(4).
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S.J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22, pp.129-146. <https://doi.org/10.1146/annurev-polisci-051117-073034>.
- Izzo, R., Smith, J. & Johnson, K. (2018). Cumulative learning in democratic systems. *Journal of Theoretical Politics*, 30(2), 145(167).
- Kalmoe, N.P. (2020). Uses and abuses of ideology in political psychology. *Political Psychology*, 41(4), pp.771-793.

Klašna, M. (2015). Corruption and the incumbency disadvantage: Theory and evidence. *Journal of Politics*, 77, pp.928–942.

Krause, S. & Mendez, F. (2009). Corruption and elections: An empirical study for a cross-section of countries. *Economics & Politics*, 21(2), pp.179-200.

Kumar Jha, C. (2023). Condoning corruption: Who votes for corrupt political parties? *Journal of Economic Behavior & Organization*, 215, pp.74-88.

Lee, I-C., Chen, E.E., Yen, N-S., Tsai, C-H. & Cheng, H-P. (2017). Are we rational or not? The exploration of voters' choices during 2016 presidential and legislative elections in Taiwan. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2017.01762>.

Levitsky, S. & Ziblatt, D. (2018). *How Democracies Die*. Crown, p.320.

Martini, S. & Quaranta, M. (2019). Political support among winners and losers: Within- and between-country effects of structure, process and performance in Europe. *European Journal of Political Research*, 58(1), pp.341–361.

McMann, K.M., Seim, B., Teorell, J. & Lindberg, S. (2017). Democracy and corruption: A global time-series analysis with V-Dem data. Working Paper Series 2017:43, Varieties of Democracy (V-Dem) Institute.

Nai, A. (2019). Voter information processing and political decision making. In: J. D. Wright (ed.) *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, Elsevier. <https://doi.org/10.1093/acrefore/9780190228637.013.912>.

Nadeau, R., Daoust, J-F. & Dassonneville, R. (2023). Winning, losing and the quality of democracy. *Political Studies*, 71(2), pp.483-500.

Nasr, M. (2021). The motivated electorate: Voter uncertainty, motivated reasoning, and ideological congruence to parties. *Electoral Studies*, 72, 102344.

North, D.C., Wallis, J.J. & Weingast, B.R. (2009). *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.

Nyhan, B. & Reifler, J. (2015). The effect of fact-checking on elites: A field experiment on U.S. state legislators. *American Journal of Political Science*, 59(3), pp.628-640. <https://doi.org/10.1111/ajps.12162>.

Pande, R. (2011). Can informed voters enforce better governance? Experiments in low-income democracies. *Annual Review of Economics*, 3(1), pp.215-237. <https://doi.org/10.1146/annurev-economics-111809-125045>.

Palmer, A. & McClendon, G. (2025). Cynicism and voter support for openly corrupt candidates: Evidence from Kenya. Working Paper, Dartmouth College & New York University.

Paulson, S. (2023). The very idea of rational irrationality. *Politics, Philosophy & Economics*, pp.1-19. <https://doi.org/10.1177/1470594X231177640>.

Rienks, H. (2023). Corruption, scandals and incompetence: Do voters care? *European Journal of Political Economy*. <https://doi.org/10.1016/j.ejpoleco.2023.102441>.

Seeberg, H.B., Slothuus, R. & Stubager, R. (2016). Do voters learn? Evidence that voters respond accurately to changes in political parties' policy positions. *West European Politics*, 40(2), pp.1-21. <https://doi.org/10.1080/01402382.2016.1245902>.

Sinclair, T.J. (2005). *The New Masters of Capital: American Bond Rating Agencies and the Politics of Creditworthiness*. Cornell University Press.

Somin, I. (2016). *Democracy and Political Ignorance: Why Smaller Government is Smarter*. Stanford University Press.

Stoetzer, L.F., Leeman, L. & Traunmueller, R. (2024). Learning from polls during electoral campaigns. *Political Behavior*, 46, pp.543-564. <https://doi.org/10.1007/s11109-022-09837-8>.

Tappin B.M., Pennycook G., and Rand D.G. (2020), Bayesian or biased? Analytic thinking and political belief updating. *Cognition* 204,104375.

Van Remoortere, A. & Vliegthart, R. (2025). Affective polarization and political (dis)trust: Investigation of their interconnection and the moderating role of (social) media use. *European Journal of Communication*, September 2025.

Vaeth, M. (2024). Rational voter learning, issue alignment and polarization. Working Paper, Paris School of Economics.

Wantchekon, L. (2003). Clientelism and voting behavior: Evidence from a field experiment in Benin. *World Politics*, 55(3), pp.399-422. <https://doi.org/10.1353/wp.2003.0010>.

Weitz-Shapiro, R. & Winters, M.S. (2017). Can citizens discern? Information credibility, political sophistication, and the punishment of corruption in Brazil. *The Journal of Politics*, 79(1), pp.60-74. <https://doi.org/10.1086/687536>.

Wood, A.K. & Grose, C.R. (2022). Campaign finance transparency affects legislators' election outcomes and behavior. *American Journal of Political Science*, 66(2), pp.516-534. <https://doi.org/10.1111/ajps.12676>.