A Study on Electoral Accountability

Alejandro Saporiti*[a] and Yizhi Wang[b]

[a]University of Manchester, alejandro.saporiti@manchester.ac.uk
[b]Tianjin University, yizhi.wang@tju.edu.cn

February 18, 2025

Abstract

We study electoral accountability by modelling and experimentally testing how parties, competing for office, simultaneously choose taxes, public-good provision, and political rents (corruption). Using a probabilistic voting framework, we derive equilibrium policies under varying institutional parameters—such as power-sharing disproportionality, transparency, and partisanship—and then we subject a subset of these predictions to a controlled laboratory test. A distinctive feature of our lab design is that both candidates and voters are human participants, enabling us to capture the strategic interplay on both sides, and disentangle the contribution of human cognitive and behavioral factors to electoral accountability. Our findings confirm that heightened electoral stakes (via a disproportional rule) deter rent-seeking and boost public goods, yet corruption remains above theoretical benchmarks. Voters do punish corruption, but partisan loyalties often blunt effective accountability.

JEL Codes: C90, D72, D73.

Keywords: Electoral accountability, corruption, institutions, voter behaviour.

1 Introduction

When citizens can hold political leaders accountable at the ballot box, public policy should, in theory, mirror the preferences of the electorate while deterring corruption

*Corresponding author.

and other self-serving behaviors. Electoral accountability is a cornerstone of democratic theory, yet empirical studies from around the world reveal a persistent gap between this ideal and actual outcomes. High levels of corruption, inefficient public service delivery, and tax policies that diverge from voter priorities all underscore the limitations of electoral competition as a solitary mechanism for aligning politicians' actions with the public interest.

Existing research highlights several factors that can undermine electoral accountability. First, voters often struggle to accurately detect or penalize corrupt behavior, especially when partisan biases and limited fiscal transparency obscure the link between observable fiscal indicators—such as tax rates and public goods provision—and underlying corruption. Second, institutional arrangements, notably power-sharing mechanisms that distribute legislative and executive authority, can dilute accountability by making it harder for citizens to pinpoint responsibility for policy outcomes. Finally, political actors are frequently incentivized to extract rents, with these incentives shaped by office salaries and the transaction costs associated with legal enforcement, penalties, and imprisonment.

This paper examines electoral accountability in a model that include taxes, public goods, and political rents (corruption) together. That is, it takes these three dimensions as interdependent components of fiscal policy. Voters observe (or infer) each candidate's policy proposals in these domains, weighing both their private costs and benefits (through taxes and consumption) and their broader social implications (public goods, corrupt behavior). By analysing voter responses to these policy bundles, the paper aims to understand how institutional arrangements—such as more or less *disproportional* power-sharing—or the extent of partisan bias can shift the equilibrium toward or away from accountable governance.

To be more specific, we develop a stylized three-stage game of policymaking, drawing on the probabilistic voting framework. In Stage 1, two political parties (or candidates) publicly commit to fiscal platforms encompassing (i) a proportional tax rate, (ii) a public good, and (iii) political rents, interpretable as corruption. Each party faces a non-empty, compact, and convex strategy set, respecting the feasibility constraint that total spending on the public good plus political rents does not exceed collected tax revenue.

In Stage 2, voters observe (or rationally infer) each party's proposed fiscal policy. Each voter derives utility from private consumption (their after-tax income), from the public good, and from the disutility of corruption. Voters also hold partisan preferences. They vote for the party whose policy proposal delivers higher expected utility, factoring in their ideological biases. An overall vote share then determines each

party's *power share* in shaping the final policy, according to a smooth contest success function that can shift from perfectly proportional to highly disproportional (winner-take-all).

Finally, in Stage 3, the implemented policy emerges as a weighted average (compromise) of the two campaign proposals, where the weights are each party's post-election power share. Both parties then derive payoffs from holding office and obtaining political rents, net of transaction costs (risk of detection or legal punishment). Voters receive their post-tax private consumption plus any benefit from the public good, minus their perceived loss from corruption.

Within this framework, we show that under mild regularity conditions, a unique pure-strategy equilibrium exists. In equilibrium, each party balances the temptation to extract political rents against the need to secure votes and thus influence over final policy. Our comparative statics highlight several drivers of accountability. First, high *power disproportionality* (where a slight majority of votes yields a large policy advantage) can push parties to reduce corruption, as the electoral stakes of losing become more severe. In turn, this greater focus on victory can lead to higher public-good provision and lower taxes, aligning policy more closely with voters' preferences.

Second, higher politician wages—or broader ego rents from holding office—raise the cost of electoral defeat, lowering equilibrium corruption levels. Similarly, when voter aversion to corruption intensifies, parties internalize a higher risk of being penalized at the polls for excessive political rents. Third, institutional or legal features that increase the transaction costs of rent-seeking (e.g., stricter enforcement or greater penalties) also deter corruption. Lastly, partisan bias can either weaken or reinforce these accountability mechanisms. A heavily polarized electorate may be less responsive to policy performance, allowing politicians to maintain rents. By contrast, when a significant fraction of voters is "swing" or ideologically neutral, parties must compete more vigorously on policy attributes, leading to lower corruption and higher provision of public goods.

Our analysis also juxtaposes the equilibrium fiscal outcomes with a benchmark *social optimum*, in which a utilitarian planner sets policy to maximize aggregate welfare. Comparing the model's equilibrium with this benchmark reveals inefficiencies across all three dimensions: higher corruption, lower public-good provision, and a tax rate that does not always reflect voters' broader interests. We demonstrate how each institutional parameter (e.g., power disproportionality, legal enforcement, partisanship, politicians' wages, etc.) affects the size of this welfare gap, emphasizing that corruption is only one piece of a more complex accountability puzzle.

To test our theoretical conclusions, we conduct a lab experiment reflecting the

structure of our election game. Each session of the experiment simulates repeated elections in small groups, with participants assigned roles as candidates or voters. Candidates propose fiscal platforms specifying a tax rate (budget), a public-good allocation that benefits all voters, and a second good that benefits only the politicians. Voters then choose between the two candidates, knowing how their income is taxed, how much of the budget goes to the public good, and—under partial or full information—how much is siphoned off as rents (second good).

A distinctive feature of our experiment is that both candidates and voters are human participants in each election, rather than focusing on just one group. This setup provides a more realistic electoral environment by capturing the simultaneous and reciprocal interactions between candidate strategies and voter decisions in real time. Unlike studies that rely on preprogrammed or hypothetical voters, our approach ensures that both campaign pledges and ballot choices arise from genuine human behavior. However, running each round in a fully human setting is more time-intensive, reducing the total number of rounds that can be implemented. Consequently, subjects have fewer opportunities to learn or refine their strategies across repeated interactions, potentially tempering the speed and depth of their convergence to theoretical benchmarks.

To compare our fully human setup against the standard approach in electoral competition experiments, where voters are computer algorithms, we also include a treatment with automated voters. Comparing this automated-voter baseline to the fully human-voter setting sheds light on how real voters' biases, heuristics, or strategic uncertainties might influence accountability outcomes. In particular, it reveals whether—and to what extent—algorithmic decision rules deviate from or align with the preferences of human voters, thereby helping us disentangle the unique contribution of human cognitive and behavioral factors to electoral accountability.

Our data confirm that electoral accountability extends far beyond the question of corruption alone. First, we find that under a disproportional rule, winning candidates propose significantly lower corruption and higher public goods—consistent with the theoretical claim that higher stakes in electoral competition deter rent-seeking. Yet, corruption remains above the model's predicted level, revealing persistent inefficiencies that underscore how real-world constraints (e.g., incomplete learning or partisan bias) keep actual outcomes from reaching the theoretical benchmark.

Second, fiscal transparency affects taxation and public-goods provision more strongly than it does corruption per se. Despite more available information, candidates still exploit some latitude for self-enrichment, suggesting that voters may not perfectly discipline politicians for corruption in practice, particularly if other policy

dimensions—such as public good levels—align with voters' preferences.

Third, the presence of automated (as opposed to human) voters amplifies both the reward for public-good provision and the penalty for corruption, suggesting that real human voters in the lab might exhibit greater partisan or behavioural biases that weaken accountability. Likewise, an increased fraction of "swing" participants heightens electoral competition, leading to tax and public-good proposals closer to the theoretical equilibrium.

Finally, fourth, candidate behavior shows systematic deviations from the model's precise numerical equilibrium: corruption is generally higher and public goods are lower than theoretically predicted. On the voter side, logistic regressions reveal that while monetary incentives matter—voters do reward lower corruption—strong partisan preferences significantly dilute these effects. This is particularly noticeable under proportional representation, where the compromise shape of the implemented policy makes the partisan bias more salient relative to the utility differences associated with the candidates' proposals. This finding confirms that ideological loyalties can shield politicians from being fully held accountable at the polls.

Taken together, these results highlight the broader nature of electoral accountability: corruption, taxes, and public goods are jointly determined, and voters' capacity—or willingness—to punish self-serving behavior shapes how well democratic processes succeed in curbing inefficiencies. By comparing the experimental outcomes with our theoretical predictions, we identify the factors that most strongly sustain accountability and those that still leave room for further policy distortions.

The rest of this paper is organized as follows. Section 2 discusses the literature more closely related to this work. Section 3 sets up the strategic framework, describing the key actors, the budget constraint, the timing of the election, and the equilibrium concept. Section 4 analyses the equilibrium of the game, providing existence, uniqueness, and comparative statics results. Section 5 details the experimental design, including treatments, procedures, and empirical strategy. Section 6 presents the empirical findings, contrasting candidate and voter behavior across different institutional and informational settings. Finally, Section 7 concludes by reflecting on policy implications, limitations of the study, and avenues for future research. For convenience of exposition, all formal proofs are collected at the end in Appendix A.

2 Related Literature

This paper contributes to the broader literature on electoral accountability and political corruption. We do not aim to survey all the existing research here, but rather

highlight closely related studies that integrate theory and, when available, experimental evidence.

Our theoretical setup builds on the probabilistic voting framework with differentiated candidates (Lindbeck and Weibull, 1987), as in Afridi et al. (2024), Besley et al. (2010), Heggedal et al. (2022), Persson and Tabellini (2000), Polo (1998), among others.[1] A key departure in our work is that we introduce a *reduced-form* post-election policymaking mechanism, in which parties' proposals are merged into a final policy according to each party's *power share*. By representing the degree of *disproportionality* in power-sharing with a single parameter, our model can capture a continuum of institutional rules that move from fully proportional influence to near winner-take-all systems. This provides a parsimonious way to analyse how institutional design affects corruption, taxation, and public-goods provision within a single framework.[2]

A prominent theme in the literature is that more competitive or majoritarian-leaning electoral environments reduce rent-seeking by heightening accountability (Chang and Golden, 2007, Ferraz and Finan, 2011, Golden and Mahdavi, 2015, Kunicová, 2006, Kunicova and Rose-Ackerman, 2005, Persson and Tabellini, 2000, Persson et al., 2003, Rudolph and Däubler, 2016). While a few studies point to exceptions or non-monotonicities in this relationship (for example, Afridi et al., 2024 note a U-shaped effect of competition on corruption), the conventional wisdom is that greater electoral stakes prompt politicians to moderate corruption in pursuit of votes. We build on this idea by allowing parties to choose "political rents" as one part of a fiscal platform that also includes taxes and public goods. By systematically manipulating *power-sharing disproportionality* in our experiment, we provide new, controlled-evidence on whether tighter competition indeed reduces corruption—and whether it interacts with voter behavior in ways not easily observable in field data.

In the empirical literature on electoral accountability, most studies focus on cross-country evidence within a single policy dimension (Foucault et al., 2017, Lowry et al., 1998, Min, 2008). By contrast, we employ a laboratory setting that allows us to observe voting behavior across three distinct fiscal policy components simultaneously and investigate how information, partisan bias, voter gender, and institutional rules influence accountability. This approach aligns with the broad literature on electoral accountability (Arias et al., 2019, Ashworth et al., 2017, Bruns and Himmler, 2016, Devdariani and Hirsch, 2023, Eggers et al., 2018, 2014, Hellwig and Samuels, 2008,

[1] Alternative modelling approaches include principal-agent or career-concerns frameworks (see Cubel et al., 2024, Martinelli, 2022, Martinelli and Esandi, 2022).

[2] In contrast, the related literature on accountability often relies on political agency models that either focus exclusively on corruption or address policies in isolation (Besley and Case, 1995, Duggan and Martinelli, 2020, Martinelli, 2022, Zudenkova, 2018).

Jones, 2010, Martin and Raffler, 2021), which underscores the importance of information availability, institutional structures, and partisan biases in shaping politicians' and voters' behavior. Yet relatively few studies employ experimental methods to directly test these predictions.

Because corruption is difficult to measure in field settings (Olken, 2009, Rose, 2014, Treisman, 2014), many researchers have turned to laboratory and field experiments. Early studies tended to focus on bribery (see surveys by Abbink, 2006, Bobkova and Egbert, 2012, Isaac and Norton, 2012), while more recent work incorporates formal electoral frameworks (e.g., Cubel et al., 2024, Heggedal et al., 2022, Martinelli and Esandi, 2022). Notably, Heggedal et al. (2022) model a probabilistic voting environment and show that although raising politicians' wages can lower corruption in noise-free settings, voter preference shocks can produce more nuanced outcomes. Our study builds on this foundation in two key ways. First, we add taxes and public goods into the electoral game to examine how corruption interacts with other fiscal decisions. Second, we introduce a continuum of *power-sharing* rules—ranging from fully proportional to highly disproportional—to offer new insights into how electoral institutions shape policy outcomes and rent extraction.

Additionally, an important methodological contribution of our experimental design is that both *candidates* and *voters* are human participants, whereas most past studies use automated agents on one side. This dual-human setup, combined with an automated-voter treatment for comparison, lets us pinpoint how real human voters differ from algorithmic decision-makers and whether such differences systematically affect accountability.

Finally, we contribute to the growing literature on partisan or group biases in corruption experiments (Anduiza et al., 2013, Cubel et al., 2024, Eggers et al., 2014, Solaz et al., 2019). Our findings suggest that partisanship may dampen electoral punishment for corruption—unless the electoral system is structured to incentivize intense competition for swing voters. Hence, we shed light on a possibly overlooked interaction between partisan polarization and institutional design.

3 The Model

This section outlines the strategic framework by introducing the key actors and their respective payoffs. It then provides a detailed description of the game, its temporal structure, and the concept of equilibrium.

Political parties. There two political entities (parties or candidates), denoted as A and B, that engage in a singular electoral contest to establish a government. The elected government is responsible for providing a public good $g \in \mathbb{R}_+$ to the economy, subject to the budget constraint $t \cdot y \geq g + r$, where $t \in [0,1]$ represents a proportional income tax rate, y signifies the total income of the economy, and $r \in \mathbb{R}_+$ corresponds to the tax revenue diverted to political rents (interpreted as corruption).

During the electoral campaign, each party $C = A, B$ independently and simultaneously commits to a policy proposal $\mathbf{x}_C = (t_C, g_C, r_C) \in X = [0,1] \times \mathbb{R}_+ \times \mathbb{R}_+$, where X is non-empty, compact, and convex, and $\mathbf{x}_C$ satisfies the budget constraint. Upon the conclusion of the election, the parties jointly determine the fiscal policy $\mathbf{x} = (t, g, r)$. We postulate that $\mathbf{x} = \rho \mathbf{x}_A + (1 - \rho)\mathbf{x}_B$, where $\rho : X \times X \to [0,1]$ symbolises party A's share of power (or influence) in the policymaking process. This influence is dictated by the relative strengths of the parties in the election, represented by their ratio of votes. In accordance with the literature on contest success functions (refer to Tullock 1980, Hirshleifer 1989, Garfinkel and Skaperdas 2007), we assume that

$$\rho = \left(1 + \left(\frac{1-s}{s}\right)^{\eta}\right)^{-1}, \tag{1}$$

where $s : X \times X \to [0,1]$ denotes party A's vote share, and $\eta \in [1, \infty)$ is a finite parameter representing the disproportionality of the power-sharing rule.[3]

The rule stipulated above does not bias the election results when $s = 0, 1/2,$ or 1, while it may disproportionately favour the votes of the winner and symmetrically disadvantage the votes of the loser otherwise. However, such distortions are smooth, and no discontinuity occurs (particularly, at $1/2$). When $\eta = 1$, the above expression represents the proportional rule, where the influence of each party aligns with its vote share. As η increases beyond 1, the rule becomes increasingly disproportionate and biased towards the party with the larger vote share. Ultimately, as η approaches infinity, (1) encapsulates the winner-take-all rule, where the party with the majority of votes assumes unilateral control of the government and sets policy.

Electorate. The economy comprises a finite set $N = \{1, 2, \ldots, n\}$ of voters. Each voter $i \in N$ possesses a pre-tax income $y_i > 0$, with $\sum_{i \in N} y_i = y$. The preferences of voter i over the policy space X are represented by a utility function $V_i(t, g, r) = u_i((1-t)y_i) + a_i(g) - \beta_i r$, where $u_i(\cdot)$ and $a_i(\cdot)$ are two strictly increasing, concave, and C^2 functions.

[3] An alternative approach would be to model each party's power share as a function of the *margin of victory* (or electoral mandate), rather than the *ratio of votes*. The qualitative findings of this paper remain robust under this alternative specification, as the equilibrium characterisation under the "margin of victory" power-sharing rule only exhibits minor alterations compared to that derived under (1).

The parameter $\beta_i \geq 0$ quantifies voter i's dis-utility from corruption, which is disliked not only because it reduces the provision of the public good, but also because voter i can consider unfair that politicians privately benefit from taxpayers' contributions. That is, $V_i(\cdot)$ is compatible with the existence of a "social motive" for condemning corruption, capturing the "selfish" case where $\beta_i = 0$ as a special instance.

In addition to policies, each voter has direct concerns about the parties. One reason is that a political party's effectiveness in implementing its policy if elected could depend on idiosyncratic characteristics that cannot be altered during the election period. Another reason is that parties might be unable to credibly change their positions on some issues, and the election is fought over the issues for which they can alter their stance. We denote $\theta_i \in \mathbb{R}$ as the partisan preference (or bias) of voter i towards party B, where θ_i is drawn from a C^2 cumulative distribution $F_i(\cdot)$, with density $f_i(\cdot)$ positive everywhere over the differences $d_i(\mathbf{x}_A, \mathbf{x}_B) = V_i(\mathbf{x}_A) - V_i(\mathbf{x}_B)$, and with $f_i(0) = f(0)$ for all $i \in N$ (i.e., everybody faces ex ante the same probability of being a partisan neutral).

Payoffs. Each voter i seeks to maximize their overall utility, which is the sum of the utility from policies $V_i(t, g, r)$ and the partisan bias θ_i. As for the parties, each strives to exert influence over policy and to seize political rents. Specifically, the payoff for party A (and similarly, party B) is given by the function $\Pi_A(\mathbf{x}_A, \mathbf{x}_B) = \rho \cdot (R + (1-\gamma) r_A)$ (and correspondingly, $\Pi_B(\mathbf{x}_A, \mathbf{x}_B) = (1-\rho) \cdot (R + (1-\gamma) r_B)$), where $R > 0$ represents the utility from holding power (e.g., politicians' salaries), and $\gamma \in [0, 1]$ is a transaction cost associated with seizing political rents (e.g., the probability of being caught and sent to prison). The rationale behind the second term, that is, $\rho(1-\gamma) r_A$ (and correspondingly, $(1-\rho)(1-\gamma) r_B$), is that since the implemented policy is a blend of the election pledges, each party benefits from corruption (net of the transaction cost) to the extent permitted by its influence in the policymaking process.

Game structure. The environment described thus far defines a sequential game with three stages:

Stage 1: Political campaign

i. Each party $C = A, B$ simultaneously and independently commits to a policy platform $\mathbf{x}_C \in X$, where $\mathbf{x}_C = (t_C, g_C, r_C)$ satisfies $g_C + r_C \leq t_C y$;

Stage 2: Electoral process

i. The partisan bias θ_i of each voter $i \in N$ is drawn from $F_i(\cdot)$ by nature;

ii. Upon observing $(\mathbf{x}_A, \mathbf{x}_B)$ and θ_i, each voter $i \in N$ votes for party A if $V_i(\mathbf{x}_A) > V_i(\mathbf{x}_B) + \theta_i$, flips a coin if $V_i(\mathbf{x}_A) = V_i(\mathbf{x}_B) + \theta_i$, and votes for B otherwise;[4]

iii. The vote share $s(\mathbf{x}_A, \mathbf{x}_B) = \frac{1}{n} \sum_{i \in N} F_i(V_i(\mathbf{x}_A) - V_i(\mathbf{x}_B))$ is determined;[5]

Stage 3: Policy implementation

i. The fiscal policy $\mathbf{x} = \rho \mathbf{x}_A + (1 - \rho)\mathbf{x}_B$ is determined by the platforms $(\mathbf{x}_A, \mathbf{x}_B)$ in conjunction with $s(\mathbf{x}_A, \mathbf{x}_B)$ and the power-sharing rule in (1);

ii. Each party C receives its payoff $\Pi_C(\mathbf{x}_A, \mathbf{x}_B)$, and every voter i obtains the utility $V_i(\mathbf{x})$, plus the partisan preference θ_i if i voted for B.

In relation to the game structure outlined above, two key points merit special attention and a few additional comments. Firstly, for the sake of simplicity, we do not explicitly model the policymaking process that unfolds after the election. Instead, we depict the final stage in a reduced form, where the fiscal policy is a linear combination of the policy proposals of the parties, each weighted by the power each party secured in the election. This approach follows the modelling strategy of Herrera et al. (2016), Matakos et al. (2016), Saporiti (2014), who represent power-sharing with the help of a contest success function. This function is meant to reflect in a reduced-form the institutions of representation, governance, and policymaking (such as, separation of powers, the electoral rules, forms of government, agenda-setting and veto powers, etc.) that shape the mechanism that transforms the votes of the parties into the decision-making power or "influence" over the implemented policy.

Secondly, at the election, each voter is presented with a binary choice: to vote for party A or party B. This stage necessitates further explanation on two fronts. First, voters make their choice after observing their partisan bias and the proposals of both candidates. With respect to the latter, the game does not differentiate between scenarios where the electorate *observes* the level of political rents of the parties, and those where voters *infer* corruption as a residual from the budget constraint upon being informed of the other components of fiscal policy. These two scenarios are indistinguishable in our theoretical analysis due to the assumption of voter rationality. That is, voters are assumed to possess the ability to analyse and interpret information effectively, with perfect capacity to deduce the level of corruption from public good spending and collected taxes. This assumption about voter behaviour will be further

[4]The partisan bias could alternatively be a multiplicative factor, implying that voter i votes for party A if $V_i(\mathbf{x}_A) > \theta_i V_i(\mathbf{x}_B)$, and so forth. Given that the logarithm of $V_i(\cdot)$ represents the same preferences as $V_i(\cdot)$, the results for the additive case extend directly to the multiplicative one.

[5]Given that $F_i(\cdot)$ is continuous, the probability that $V_i(\mathbf{x}_A) = V_i(\mathbf{x}_B) + \theta_i$ has measure zero. Thus, the probability that voter i votes for party A can be associated with the event $V_i(\mathbf{x}_A) \geq V_i(\mathbf{x}_B) + \theta_i$.

scrutinised in the empirical section, where subjects will be exposed to varying degrees of publicly available information regarding corruption.

Second, given any expectation about the electoral outcome and any rule that determines the ex-post implemented policy, it is optimal for every voter to support the party whose *platform* promises the highest utility.[6] By anticipating voters' platform-based decisions, both parties converge to the same policy in equilibrium—that is, they propose identical platforms. Consequently, once votes are cast and the ex-post implemented policy is formed (often as a weighted average of the party platforms), it coincides with either platform. From the voter's perspective, there is no meaningful distinction between "voting for a platform" and "voting for the implemented policy." At least in symmetric equilibria, therefore, the ex-ante expectation that "my chosen platform will shape the ultimate policy" is indeed confirmed ex post.

One can also reverse the argument and start by assuming that the parties converge on a single policy in a symmetric equilibrium. Under that assumption, voters would be indifferent between choosing a platform and choosing the ex-post policy, since both ultimately yield the same outcome. Either way—whether one begins by positing platform-based voting or a symmetric equilibrium—the conclusion is the same: in a symmetric equilibrium, the ex ante (platform-based) and ex post (implemented-policy) perspectives are effectively identical.

This observation simplifies the strategic analysis of the three-stage election game to the analysis of an election game in which the equilibrium values of the policy variables are determined by the parties in the initial stage, with full anticipation of the electorate's optimal (Nash equilibrium) strategies in the subsequent stage. We denote this election game as $\mathcal{G} = (X, \Pi_C)_{C=A,B}$.

Equilibrium. An equilibrium of $\mathcal{G} = (X, \Pi_C)_{C=A,B}$ is defined as a profile $(\hat{\mathbf{x}}_A, \hat{\mathbf{x}}_B) \in X \times X$ such that for all $C = A, B$, and all $\mathbf{x}' \in X$, $\Pi_C(\hat{\mathbf{x}}_C, \hat{\mathbf{x}}_{-C}) \geq \Pi_C(\mathbf{x}', \hat{\mathbf{x}}_{-C})$.

4 Equilibrium Analysis

To start our discussion, we lay down a condition that guarantees the existence of a unique equilibrium in pure strategies for the election game $\mathcal{G}$. This condition imposes constraints on the cumulative distributions $F_i(d_i(\cdot))$ and the power share rule $\rho(s(\cdot))$. To be precise, fix an arbitrary party C and the policy $\mathbf{x}_{-C} \in X$ of the opposing party. Consider next any composite function $\psi_C(\cdot)$ of $v_C(\cdot, \mathbf{x}_{-C}))$ on X, both of which are twice continuously differentiable. For each $\mathbf{x}_C \in X$, we define $\delta_{ij}(\mathbf{x}_C) =$

[6] See, for example, Saporiti (2014), Matakos et al. (2016), and Bol et al. (2019) for similar reasoning.

$\frac{\partial^2 v_C(\mathbf{x}_C,\mathbf{x}_{-C})/\partial x_i\,\partial x_j}{\partial v_C(\mathbf{x}_C,\mathbf{x}_{-C})/\partial x_i \cdot \partial v_C(\mathbf{x}_C,\mathbf{x}_{-C})/\partial x_j}$, with $x_i, x_j \in \{t,g,r\}$, and let $\delta(\mathbf{x}_C) = \frac{\psi_C''(v_C(\mathbf{x}_C,\mathbf{x}_{-C}))}{\psi_C'(v_C(\mathbf{x}_C,\mathbf{x}_{-C}))}$. We then construct for each $\mathbf{x}_C \in X$ the matrix $\Delta(\mathbf{x}_C) = \left(\delta_{ij}(\mathbf{x}_C)\right)_{x_i,x_j\in\{t,g,r\}}$, with $\Delta^k(\mathbf{x}_C)$, $k = 1,2,3$, signifying the determinant of the k-th leading principal minor of $\Delta(\cdot)$ at $\mathbf{x}_C$. Lastly, given any square matrix $M = \left(m_{ij}\right)$, we define $M_{[i,u]}$ as the matrix derived from M by substituting in each column of the i-th row the entry scalar u.

Our sufficient condition for the existence of an equilibrium is as follows:

Condition ℂ. For each political party $C = A, B$, each voter $i \in N$, and each policy $\mathbf{x}_{-C} \in X$, the function $F_i(d_i(\cdot,\mathbf{x}_{-C}))$ (and correspondingly, $\rho(s(\cdot,\mathbf{x}_{-C}))$) satisfies that for every $k = 1, 2, 3$, and all $\mathbf{x}_C \in X$, $(-1)^{k+1}\,\Delta^k(\mathbf{x}_C) < (-1)^k \sum_{j\le k} \Delta^k_{[j,\,\delta(\mathbf{x}_C)]}(\mathbf{x}_C)$.

Condition ℂ ensures that when voters' preferences are mapped into parties' payoffs, the resulting function retains its concavity. This transformation from voters' utilities into parties' payoffs unfolds in two stages. Initially, voters' preferences are converted into a vote share for each party via the distribution of the partisan biases. Subsequently, each vote share is translated into a power share through the power-sharing rule specified in (1).

Condition ℂ stipulates that the probability distribution $F_i(d_i(\cdot))$ for each voter i should not be excessively convex, thereby allowing the vote share function to inherit the concavity of the individual utility functions. Probability distributions that satisfy this condition include the uniform distribution and the doubly exponential distribution (logit) function as considered by Lindbeck and Weibull (1987).

Furthermore, condition ℂ sets an upper limit for the parameter η of power-sharing disproportionality. The power-sharing rule (1) exhibits an increasing trend between 0 and 1, but it assumes a concave shape when the vote share exceeds $1/2$, and a convex shape when it falls below $1/2$. Consequently, the constraint on η is designed to ensure that the convexity across the range of vote share values between 0 and $1/2$ is not overly pronounced, thus preserving the concavity of the vote share when it is transformed into the power share.

Having established condition ℂ, we are now in a position to present our findings on the existence and uniqueness of equilibrium.

Proposition 1 (Equilibrium Existence)**.** *If condition* ℂ *is satisfied, then the election game* $\mathcal{G}$ *has a unique equilibrium.*

The validation of this proposition leverages the Debreu-Glickersberg-Fan's existence theorem for continuous and quasi-concave games. Given that each party's strategy set is non-empty, convex, and compact, and its payoff function is continuous, the

proof hinges on demonstrating that the conditional payoff function exhibits (strict) quasi-concavity on the party's own strategy. This is ensured by invoking condition ℂ. For the sake of clarity, the technical details are elaborated in Appendix A.

While Proposition 1 is a crucial component of the analysis, establishing the existence and uniqueness of the equilibrium is merely the beginning. To harness the predictive power of the model, it is essential to delineate the properties of the strategies deployed at equilibrium. The next proposition addresses this aspect, laying a solid groundwork for the empirical exploration in the subsequent section. It sheds light on the complex interplay between political rents (and correspondingly, the provision of public goods) and a host of other factors. These encompass the disproportionality in power-sharing, the remuneration received by politicians, partisan preferences, and the transaction costs and intrinsic disutility associated with political corruption.

Proposition 2 (Equilibrium Characterization)**.** *Let* $(\mathbf{x}_A^*, \mathbf{x}_B^*) \in int(X \times X)$ *be the equilibrium of the game* $\mathcal{G}$*. For each party* $C = A, B$*, the policy* $\mathbf{x}_C^* = (t^*, g^*, r^*)$ *is determined by* $t^* = \frac{r^*+g^*}{y}$ *and equations (2)–(3):*

$$((1-\gamma)\, r^* + R)\, \eta \frac{f(0)}{n} \sum_{i \in N} \left(a_i'(g^*) - u_i'((1-t^*)\, y_i) \frac{y_i}{y} \right) = 0, \tag{2}$$

$$\frac{1-\gamma}{2} - ((1-\gamma)\, r^* + R)\, \eta \frac{f(0)}{n} \sum_{i \in N} \left[u_i'((1-t^*)\, y_i) \frac{y_i}{y} + \beta_i \right] = 0. \tag{3}$$

The equilibrium, as outlined in Proposition 2, focuses on scenarios where corruption is strictly positive (i.e., on the interior of X). This aligns with our primary objective of deriving an exhaustive set of comparative statics results to inform the experimental section.[7] The ensuing proposition succinctly synthesizes the obtained results, offering a comprehensive overview of how various institutional and behavioural factors influence corruption, taxes, and public goods provision. This proposition not only provides fresh insights into these relationships but also enriches the broader discourse on corruption dynamics within diverse institutional and behavioural frameworks.

Proposition 3 (Comparative Statics—Policy)**.** *Let* $\mathbf{x}_C^* = (t^*, g^*, r^*) \in int(X)$ *be the equilibrium policy of party* $C = A, B$ *in the election game* $\mathcal{G}$*. Then: (i)* $\frac{\partial g^*}{\partial \mu} > 0$*,* $\frac{\partial r^*}{\partial \mu} < 0$*, and* $\frac{\partial t^*}{\partial \mu} < 0$ *for all* $\mu \in \{\eta, R, \gamma, \beta_i, f(0)\}$*.*

The comparative statics conducted above demonstrate that political rents are inversely related to power-sharing disproportionality (η). Specifically, a more disproportional power-sharing rule results in a reduction of corruption. This outcome can

[7] It is noteworthy however that for high values of η, the equilibrium level of corruption can be zero, as illustrated in Figure 1.

be attributed to the fact that in a system with greater disproportionality, the incentives for political parties to accumulate political rents are effectively curtailed by the competing incentive to secure electoral victory. Consequently, an increase in disproportionality, or a decrease in power-sharing, enhances the provision of public goods while simultaneously reducing the tax rate. This, in turn, improves the welfare of the electorate (refer to Proposition 5).

A similar effect is induced by the transaction cost parameter (γ), which increases the cost associated with the appropriation of political rents, thereby deterring politicians from engaging in corrupt practices. Additionally, the impact of politicians' wages (R) and voters' aversion to corruption (β_i) on political rents is also negative. The rationale is as follows: an increase in wages raises the potential cost of losing an election, incentivising politicians to maintain ethical conduct to continue benefiting from holding office. Similarly, an increase in voters' disdain for corruption heightens the punitive measures against political rent-seeking, motivating politicians to reduce their engagement in such activities.

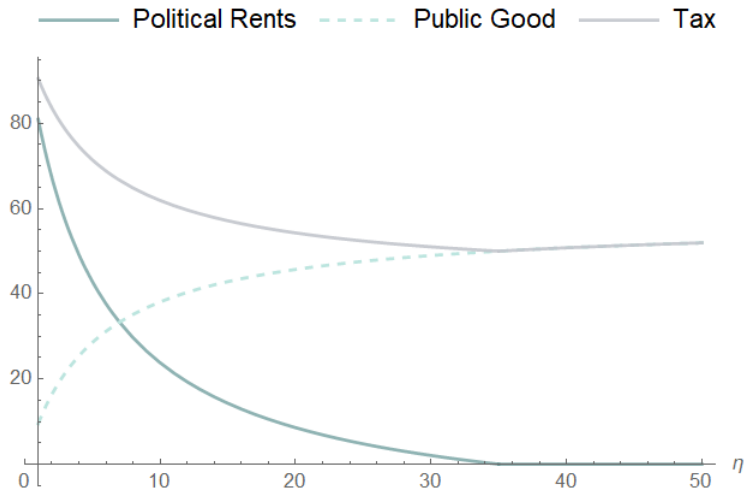
Regarding the effect of partisan preferences ($f(0)$), which can be theoretically ambiguous, the sharp prediction displayed in Proposition 3 is driven by the fact that ex ante, that is, before nature chooses the partisan biases, the probability of each voter of being biased towards neither party is positive and the same for everybody (recall we assumed $f_i(0) = f(0) > 0$ for all $i \in N$). Under this condition on voters' densities, the model demonstrates that a larger fraction of ideologically neutral (swing) voters can exercise more effective electoral control over politicians, thereby resulting in a lower level of political corruption. This result will be further scrutinised in the following section using data from the experiments.

The two previous propositions are illustrated in the following example.

Example 1. Consider a scenario where for each voter $i \in N$, $u_i(\cdot) = a_i(\cdot) = \ln(\cdot)$ and $\beta_i = 0$, and let θ_i be uniformly distributed. The equilibrium policies that emerge from equations (2) and (3) can be expressed as follows: $r^* = C^{-1} \cdot [(1-\gamma)y - 4R\eta f(0)]$, $g^* = C^{-1} \cdot [2\eta f(0)(R+(1-\gamma)y)]$, and $t^* = C^{-1} \cdot [(1-\gamma)(1+2\eta f(0)) - 2R\eta f(0)y^{-1}]$, where $C = (1-\gamma)(1+4\eta f(0))$. These policies are graphically illustrated in Figure 1. Using the closed-form solution, it can be verified that the first-order partial derivatives of corruption (and respectively, the tax rate and public good) with respect to the model parameters all exhibit the correct sign as predicted by Proposition 3.

In our final analysis in this section, we aim to quantify the welfare loss engendered by the political process. To achieve this, we derive the optimal policy, denoted as $\hat{\mathbf{x}} = (\hat{t}, \hat{g}, \hat{r})$, that a utilitarian social planner would adopt. This planner maximizes the social welfare $W(t,g,r) = \sum_i V_i(t,g,r)$, while adhering to the budget constraint.

Figure 1: Comparative Statics



Note: The graph shows in the vertical axis the equilibrium level of political rents, public good, and taxes for $n = 11$, $y = 100$, $f(0) = 0.05$, $R = 10$, and $\gamma = 0.3$. The parameter of power-sharing disproportionality is measured in the horizontal axis, and it ranges from 1 to 50.

We can describe the social optimum in the following proposition:

Proposition 4 (Social Optimum)**.** *The optimal policy* $\hat{\mathbf{x}} = (\hat{t}, \hat{g}, \hat{r})$ *dictates that political rents* $\hat{r} = 0$*, the tax rate* $\hat{t} = \frac{\hat{g}}{y}$*, and the public good* $\hat{g}$ *is determined by equation (4):*

$$\sum_{i \in N} a_i'(\hat{g}) = \sum_{i \in N} u_i'((1 - \hat{t})\, y_i)\, \frac{y_i}{y}. \tag{4}$$

Remarkably, equation (4) exhibits a notable resemblance to Samuelson's condition for the optimal provision of public goods. Essentially, (4) suggests that the aggregate of each individual's marginal benefits from the public good must balance out with its societal marginal cost. This cost is influenced by the individual marginal disutility of taxation, adjusted by each taxpayer's relative income level. This equilibrium arises from financing public goods through a proportional income tax rate, where taxation is viewed as a public bad that diminishes private consumption.

When the optimal policy $\hat{\mathbf{x}} = (\hat{t}, \hat{g}, \hat{r})$ is juxtaposed with the equilibrium $\mathbf{x}^* = (t^*, g^*, r^*)$ obtained in Proposition 2, the inefficiencies become apparent. These inefficiencies are attributable to the swing voter densities $f(0)$, which encapsulate the electoral uncertainty and the resultant lack of fierce electoral competition, thereby allowing political rents to persist. Moreover, the efficiency loss is also tied to the power-sharing rule and the cost γ associated with rent-seeking. This relationship is further illustrated in the subsequent example, which continues the analysis of Example 1.

Example 2 (Cont.)**.** The social optimum in Example 1 is given by: $\hat{r} = 0$, $\hat{g} = \frac{y}{2}$, and $\hat{t} = \frac{1}{2}$. Consequently, the gap $|\hat{\mathbf{x}} - \mathbf{x}^*|$ between the social optimum and the political equilibrium, which quantifies the inefficiencies inherent in the electoral process, is determined by $|\hat{r}-r^*| = C^{-1}\cdot|(1-\gamma)y-4R\eta f(0)|$, $|\hat{g}-g^*| = C^{-1}\cdot|y\,C/2-2\eta f(0)(R+(1-\gamma)y)|$, and $|\hat{t}-t^*| = C^{-1}\cdot|C/2-(1-\gamma)(1+2\eta f(0))+2R\eta f(0)y^{-1}|$, where $C = (1-\gamma)(1+4\eta f(0))$. Moreover, if political rents are positive, i.e., if $(1-\gamma)y > 4\eta f(0)R$, straightforward calculations show that $\frac{\partial|\hat{r}-r^*|}{\partial\mu} < 0$, $\frac{\partial|\hat{g}-g^*|}{\partial\mu} < 0$, and $\frac{\partial|\hat{t}-t^*|}{\partial\mu} < 0$ for all $\mu \in \{\eta, R, \gamma, f(0)\}$.

Drawing upon the comparative statics presented in Proposition 3, it is plausible to anticipate that a decrease in power-sharing, and respectively, an increase in either politicians' wages, corruption aversion, the cost of rent-seeking, or the fraction of independent voters, will benefit the electorate. Specifically, these changes will mitigate the welfare loss induced by the electoral process compared to the social optimum. To formally articulate this intuition, define the welfare loss as $L(\hat{\mathbf{x}}, \mathbf{x}^*) = W(\hat{\mathbf{x}}) - W(\mathbf{x}^*)$, where $\hat{\mathbf{x}}$ and $\mathbf{x}^*$ represent the utilitarian social planner's solution and the political equilibrium, respectively. The subsequent statement is derived from the results explicitly delineated in the preceding propositions.

Proposition 5 (Comparative Statics—Welfare)**.** *The welfare loss is determined by equation (5):*

$$L(\hat{\mathbf{x}}, \mathbf{x}^*) = \sum_{i\in N} u_i((1-\hat{t})y_i) + \sum_{i\in N} a_i(\hat{g}) - \sum_{i\in N} u_i((1-t^*)y_i) - \sum_{i\in N} a_i(g^*) + \sum_{i\in N} \beta_i\, r^*, \qquad (5)$$

and it is decreasing with respect to the model parameters, as indicated by the mathematical expression $\frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial\mu} < 0$ *for all* $\mu \in \{\eta, R, \gamma, \beta_i, f(0)\}$.

5 Experiment

After calculating the Nash equilibrium policies, including the corruption levels chosen by political candidates, we now turn to the experimental part of the paper. This section describes the empirical tests ran on various theoretical predictions, using a controlled experimental setup that replicates the conditions described in Section 3.

5.1 Treatments

The experiments were conducted at the laboratory facilities of the Ma Yinchu School of Economics, Tianjin University, during the fall of 2024. Table 1 outlines the treat-

ments along with their associated parameters and equilibrium values.[8] Among the various hypotheses derived from the theoretical model, we focused on testing four key hypotheses. This decision was guided by both practical and theoretical considerations.

Table 1: Parameters and Equilibrium Values for Different Treatments

Parameters/Treatments	**T1**	**T2**	**T3**	**T4**
Power Sharing (η)	1	1	1	3
Voter Behavior	Subjects	Automated	Subjects	Subjects
Fiscal Transparency	Partial	Partial	Full	Partial
Equilibrium values				
Tax rate	52.52	52.52	52.52	60.60
Public good	48.48	48.48	48.48	59.22
Corruption	4.04	4.04	4.04	1.38

Notes: The other key parameters used to compute the equilibrium policies are set as follows: ego rents $R = 0$, transaction costs $\gamma = 0$, voter aversion to corruption $\beta = 0$, and partisan bias at $f(0) = 2$.

The treatments were designed to address distinct dimensions of the model: the relationship between electoral outcomes and power-sharing, the role of transparency in fiscal policy, and the implications of voter behavior. Specifically, the first treatment implemented a proportional power-sharing rule, where policy-making power was directly proportional to the vote shares received by each candidate.

The second treatment replicated the proportional rule of the first treatment but replaced human voters with automated algorithms, aligning with the methodology commonly used in the experimental literature on electoral competition. This allowed for a closer comparison with existing studies, which often focuses exclusively on candidates' decisions, while isolating the effects of automated versus human decision-making on electoral outcomes.

The third treatment, like the first, used a proportional rule but added full transparency by explicitly informing voters about all elements of fiscal policy, including allocations to both public goods and rents. This treatment was designed to investigate how transparency affects voter accountability and candidate behavior. In the other three treatments only partial information about taxes and public good was released to the electorate in each round, leaving voters to infer corruption levels from the budget constraint.

[8]To streamline both the equilibrium calculations and the information provided to participants, we set ego rents (R), transaction costs (γ), and voter aversion to corruption (β) all to zero. This simplification was necessary given the already high complexity of the experimental environment, which demands significant attention and focus from the participants.

Finally, the fourth treatment was identical in most respects to the first, but with a key modification: policy-making power responded more than proportionally to vote shares, particularly for the candidate receiving a majority of votes. This allowed us to examine how changes in the payoff structure influenced candidates' strategic behavior and voters' decisions.

The selection of these four treatments strikes a balance between theoretical relevance and empirical feasibility. While the theoretical model generates a wide array of hypotheses, these four treatments were chosen to focus on the most salient aspects of the research exercise and to address gaps in the existing literature. They capture critical mechanisms such as power allocation, transparency, and voter behavior, all while ensuring the experimental design remains practical in terms of complexity and session length.

5.2 Experimental design

Each experimental treatment consisted of two independent sessions, each lasting 25 periods. Each period simulated an election involving two candidates (labelled A and B) and five voters (denoted as 1, 2, 3, 4, and 5). While longer sessions with 50 or 60 periods are common in experimental studies on elections, we opted for shorter sessions due to the inclusion of human voters instead of computer algorithms.

Human voters increased the time required for each election, as candidates made their decisions first, followed by the voting process, during which all five voters cast their votes independently. To manage session length while ensuring robust data collection, we maintained the smallest meaningful number of voters per election. This constraint was relaxed in treatment 2, where voters were replaced by computer algorithms, allowing a more streamlined process focused on candidates' behavior.

At the start of each session, 28 participants were randomly and anonymously assigned the role of either "candidate" or "voter." These roles remained fixed throughout all 25 elections. Participants were informed of their own roles but not the identities of others they were matched with. In treatment 2, the number of participants was reduced, as voters were automated, and only 8 human subjects were assigned the roles of candidates A and B (see Table 2 for further details).

Within each election, participants were randomly assigned to one of four groups, each consisting of seven members. Although group composition changed between elections, the initial role (candidate or voter) of each participant remained fixed. For instance, a participant assigned the role of candidate A in group 1 during the first election might be candidate A in group 3 in the second election, and so forth. The same applied to voters, except in treatment 2, where the electorate was automated.

Group re-matching was implemented in each period to mitigate potential collusion among candidates and voters.

The experimental design employed neutral language to avoid introducing unintended positive or negative connotations unrelated to the incentive structure. For example, participants assigned the role of candidates were instructed to make two decisions in each election: (a) a **budget**, representing a percentage (a number between 1 and 99) of the total voter income, set at 100 Experimental Currency Units (ECUs); and (b) an **allocation** of the collected income between **good 1**, which benefits all voters equally, and **good 2**, which benefits only the candidates.

Voters observed the budget and proposed spending on good 1 from each candidate before independently voting for their preferred candidate (A or B). In treatment 3, voters were also explicitly informed about the spending on good 2. This variation in the information setting enabled an assessment of how voters respond differently to scenarios where fiscal misuse is explicitly disclosed compared to cases where they infer such misuse indirectly from other fiscal policy components. This comparison highlights the role of transparency in fostering accountability and the challenges voters face in less transparent settings.

Participants were informed about their payoff structure at the beginning of the session. Candidates' payoffs depended on the number of votes their proposed budget and allocation received, which determined their "share of power." Power directly influenced their earnings from the amount allocated to good 2. In most treatments, power was proportional to the number of votes received. For instance, if a candidate received three votes out of five, their power share was 0.6. However, in treatment 4, power increased disproportionately with the percentage of votes received, allowing us to examine how non-linear returns affect candidates' strategic decisions and competition for voter support.

Voters' payoffs were determined by their net income, the benefit derived from good 1, and their candidate bias. Initial incomes were randomly assigned, with voters 1–4 receiving values between 18 and 25 ECUs, and voter 5 receiving the remainder of the group's total income (100 ECUs). A candidate's proposed budget affected voters' net incomes. For instance, if a voter's initial income was 25 ECUs and a candidate proposed a budget of 32%, the voter's net income would be $(100-32)\%\times 25 = 17$ ECUs. Candidate bias, a randomly assigned value between $[-0.25, 0.25]$, influenced voters' payoffs by favouring one candidate over the other, independent of policy proposals.

To assist participants in evaluating their decisions, an 'expected payoff calculator' was provided on the decision screen. This tool allowed subjects to experiment with different combinations of their own and others' choices to observe the resulting per-

period payoffs. Participants could use the calculator as many times as needed before making their decisions. A log-sheet was also provided for participants to track their earnings throughout the experiment, helping them better understand how different decisions impacted their payoffs in previous rounds.

Control questions were administered at the start of each session to ensure participants understood the instructions. After the final election period, and before receiving payment, participants completed a post-experiment survey. The survey included questions about their experience during the experiment, their decision-making process, and related factors. Responses were anonymized and did not affect participants' per-period payments.[9]

5.3 Overview of the sessions

Table 2 summarizes the main features of the laboratory sessions for each treatment, including the total number of sessions, participants, grouping structure, average payment, and overall session length.

Table 2: Overview of the Lab Experiment

Treatment	Sessions	Subjects	Groups	Elections per Session	Exchange rate	Average Payment (RMB)	Duration (mins)
T1: Prop. Rule	2	28 (8, 20)	4	25	100 ECU = 8 RMB	44.52	92
T2: Auto. Voters	2	8 (8, 0)	4	25	100 ECU = 8 RMB	41.67	54
T3: Transp.	2	28 (8, 20)	4	25	100 ECU = 8 RMB	44.57	90
T4: Disprop. Rule	2	28 (8, 20)	4	25	100 ECU = 8 RMB	42.14	83

Each treatment was conducted in two sessions, with the total number of participants per session shown in the 'Subjects' column. The numbers in parentheses indicate how many were assigned to the roles of candidates and voters, respectively. As shown, the Proportional Rule, Disproportional Rule, and Transparency treatments each involved 28 subjects split into 8 candidates and 20 voters per session, whereas the Automated Voters treatment featured only 8 subjects per session, all of whom served as candidates, with voters replaced by computer algorithms.

Each session comprised 25 election rounds, and participants received monetary compensation based on their in-game earnings, converted at the indicated exchange rate—namely, 1 ECU = 0.08 Renminbi (RMB). The 'Average Payment' column reflects the final compensation received by participants in each treatment, including any show-up fee and incentive-driven payoffs. The distinction between candidates' and voters' payoffs is omitted in the Table, but the differences were not significant. The

[9] A copy of the lab instructions, the control questions, and the post-experiment survey are available from the authors upon request.

'Duration' column includes time for instructions, control questions, decision-making in each round, and completion of the post-experiment survey, with the total session length typically ranging from 50 to 100 minutes.

5.4 Empirical strategy

Our experiment deliberately includes both human voters and human candidates in each election, rather than substituting one side with computer algorithms or focusing on only one role. This design choice enables us to capture the full electoral dynamic in the lab, more closely mirroring real-world interplay between voter behavior and candidate strategies. Because both voters and candidates make decisions in each period, we necessarily feature fewer rounds per session (25) to keep session length manageable, yet involve more participants per treatment (approximately 60) than is typical, yielding a much richer subject pool.

In each treatment, every period involves 8 groups (4 per session), each consisting of 7 participants. Over 25 periods, this setup yields *200 elections* per treatment (8×25), covering both candidates' policy choices and the votes of all 5 voters in each group. Although these 200 elections are independent *across* the two sessions (since each session recruits a separate subject pool), they are not strictly independent *within* a single session, because the same set of participants interact over 25 periods (and may be rematched in different groups). In our statistical analysis, we therefore account for potential within-group correlations, by clustering standard errors at the group level.

Because each group repeatedly interacts over multiple periods, we employ a *panel data* approach rather than aggregating across groups or periods. Aggregation would severely reduce our sample size, discarding the within-group/time variation that arises as both candidates and voters adapt their strategies over 25 rounds. By retaining data at the group-period level, we exploit the experiment's repeated-measures nature and preserve statistical power. In practice, we use random-fixed effects specifications with clustered standard errors, rather than group fixed effects, to avoid collinearity between group intercepts and treatment indicators (since each group belongs to exactly one treatment). Depending on our hypothesis, we also include period (round) controls to capture potential learning or convergence effects over time.

To align with the theoretical predictions, we conducted two main empirical analyses. First, we examined *candidate behavior* by focusing on the policy outcomes chosen by the most-voted ('winning') candidate in each group-period. We pooled the data for T1–T3, as they share similar equilibrium predictions, and estimated random-effects GLS regressions with clustered standard errors (at the group level) to account for repeated observations within groups and across rounds. Our baseline specifica-

tions included treatment dummies (e.g., T2, T3) relative to T1, the round number, and demographic controls such as the winning candidate's gender and the group's gender composition. We then compared T1 directly with T4, which has a distinct equilibrium, using the same regression framework but isolating the $T4_g$ effect.

Second, we analyzed voting behavior at the individual level, considering up to five voter decisions (candidate A or B) per group-period. This generated a rich panel dataset of voter choices, which we pooled by subsets of treatments that share the same theoretical benchmark (i.e., T1–T3 together, and T1 and T4 separately). We ran logistic regressions of Vote(A) on policy differences—such as corruption, taxes, and public good provision—along with measures of utility differences, partisan bias, and treatment-by-policy interactions. Round fixed effects are included to capture time trends or learning effects, as well as other voter-level control variables (such as initial income, gender, etc.), with standard errors clustered at the group level. This setup allowed us to test whether voters penalize corruption, reward public goods, or exhibit partisan loyalty regardless of policy, and how these patterns differ across treatments.

Taken together, these two strands of analysis—candidate behavior and voter behavior—offer complementary insights into how each institutional variation (automation, transparency, disproportional power-sharing) shifts the equilibrium policies and the electorate's response. By leveraging repeated-measures data and focusing on treatment comparisons that match the model's equilibrium predictions, we gain a clearer view of how our theoretical mechanisms play out in practice.

6 Empirical Results

To better interpret the empirical evidence generated by the experimental treatments described in Section 5, we organize our data analysis into three parts: Section 6.1 examines candidate behavior, while Sections 6.2 and 6.3 focus on voter behavior. However, before delving into these detailed analyses, Table 3 provides a concise overview of our key findings, summarizing the main empirical results and their alignment (✓) or misalignment (✗) with our theoretical predictions. This summary serves as a roadmap for understanding how each outcome variable relates to the expectations established by our model and the equilibrium analysis presented in Sections 3 and 4.

6.1 Candidate behavior

Table 4 presents the results of Random-Effects Generalized Least Squares (GLS) regressions analysing the fiscal policies—Corruption, Taxes, and Public Good—chosen

Table 3: Empirical Findings and Their Alignment with the Theory

	Empirical Finding	Theory	
Corruption	T2 and T3 show no significant effect	✓	
	T4 significantly reduces corruption over time	✓	
	Remains persistently high, above predicted value		✗
	Strong deterrent; voters punish higher corruption	✓	
Taxes	Higher taxes reduce voter support, notably in T2	✓	
	Slightly above equilibrium in some treatments		✗
	Increase under aut. voters (T2) and transparency (T3)		✗
Public Goods	Provision is significantly rewarded by voters	✓	
	Significant increase under T4	✓	
	Below equilibrium but increases over time		✗
Partisan Effects	Strong negative effect on voting for candidate A	✓	
Institutional Effects	T4 (disprop. power-sharing) does not increase voter sensitivity to fiscal policies, but mitigates partisan effects	✓	
	T3 (transparency) shows no impact on voter behavior	✓	
Behavioral Effects	Voters favour parties offering higher monetary benefits	✓	
	Automated (fully rational) voters are more responsive to fiscal incentives than human voters		✗
	Winner's gender shows no impact on policy or voting	✓	
	More females in the group (election) is associated in some cases with higher corruption and taxes		✗

by the winning candidates across Treatments T1, T2, and T3. Columns (1), (3), and (5) include all three treatments without additional controls, while Columns (2), (4), and (6) restrict the sample to Treatments T1 and T3 only, incorporating two gender control variables:[10]

- Gender of the winning candidate, coded as 1 if the most-voted candidate in the election is female, and 0 otherwise; and
- Group's gender composition, which denotes the number of females in the group.[11]

Each dependent variable—Corruption, Taxes, Public Good—is analysed in sepa-

[10]When controlling for gender in the analysis, we include only data from human participants whose gender composition is relevant to the dynamics being studied. That is, we exclude Treatment T2 from these control variables, as voters in this treatment are represented by automated algorithms, making gender composition inapplicable to the group interactions and outcomes. This distinction ensures the analysis accurately reflects the influence of gender where it is meaningful and avoids conflating results with data where gender is irrelevant.

[11]There are 244 females (61%) among the 400 winning candidates corresponding to T1 and T3. Among these observations, 202 groups (50.5%) have a male majority, that is, fewer than 4 females, whereas the remaining 198 groups (49.5%) have more than 4 women in each of them.

rate regressions across these columns. The regression specification employed is

$$\text{Fiscal Policy}_{g,t} = \alpha + u_g + \beta_2 \cdot \text{T2}_g + \beta_3 \cdot \text{T3}_g + \gamma_1 \cdot \text{Round}_t + \gamma_2 \cdot \text{Controls}_{g,t} + \varepsilon_{g,t}, \quad (6)$$

where $\text{Fiscal Policy}_{gt}$ denotes the observed fiscal policy implemented by the most-voted candidate in group g at period t, the binary treatment variables (T2_g and T3_g) are time-invariant indicators for experimental treatments (i.e., $\text{T2}_g = 1$ corresponds to the Automated Voter treatment, and $\text{T3}_g = 1$ to the Full Transparency treatment, with the baseline being the Proportional Rule Treatment T1),[12] Round_t is a group-invariant variable capturing the round number in the experiment to control for time trends or learning effects, the error term $\varepsilon_{g,t}$ represents unobserved factors influencing fiscal policy outcomes (assumed to have a mean of zero and be independently distributed), and the group-level clustering is accounted for by the term u_g.[13]

The key findings from the data are described bellow.

Corruption. In the regression models where Corruption is the dependent variable, the coefficient for Round is slightly negative in both models but remains not statistically significant. This suggests that the progression of rounds does not significantly influence the corruption level chosen by the winning candidates.

For Treatment T2 (Automated Voters), the coefficient is negative but not significant ($p = 0.301$), indicating that automated voters do not significantly alter the corruption levels compared to the baseline Treatment T1. Similarly, Treatment T3 (Full Transparency) shows a positive coefficient that is also not significant ($p = 0.325$), suggesting that releasing information about political rents on top of taxes and public good provision does not have a significant impact on corruption compared to T1.

The constant term in both models is highly significant ($p < 0.01$) and represents the observed corruption level in Treatment T1. When compared with the Nash Equilibrium level, the empirical findings suggest that the baseline corruption levels around 21.993 (without controls) and 20.504 (with controls) are substantially higher than the equilibrium prediction of 4.04 ECUs. This indicates that candidates in the experiment choose much higher corruption levels than theoretically expected.

The variable Gender of Winning Candidate does not have a significant effect on corruption. However, the other control, namely, Group's Gender Composition, has a

[12] Since our sample consists of groups from different treatments, even though the treatment variable is invariant over time within a group, it varies across groups. That's why indexing it by g makes sense: it reminds the reader that the treatment assignment is a group-level attribute. Notice, however, that while the treatment indicator is indexed by g, it is constant for all periods t within that group.

[13] Group fixed effects are excluded to avoid collinearity with single-treatment groups, but standard errors are clustered at the group level.

Table 4: Winning Candidate's Fiscal Policy (Treatments T1–T3 Pooled)

	Corruption$_{gt}$		Taxes$_{gt}$		Public Good$_{gt}$	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Round$_t$	-0.032	-0.019	0.334***	0.427***	0.367***	0.445***
	(0.0354)	(0.0493)	(0.0715)	(0.0861)	(0.0747)	(0.0882)
T2$_g$	-1.030		3.525**		4.555*	
	(0.9953)		(1.7212)		(2.5493)	
T3$_g$	0.705	0.864	2.710**	3.211***	2.005	2.327*
	(0.7161)	(0.6293)	(1.3699)	(1.1655)	(1.3371)	(1.2961)
Winner Gender$_{gt}$		-0.809		-1.676*		-0.927
		(0.7628)		(0.9924)		(1.2193)
Group Gender$_{gt}$		0.401**		1.290**		0.833
		(0.1931)		(0.5449)		(0.5143)
Constant	21.993***	20.504***	58.137***	52.148***	36.145***	31.932***
	(0.5106)	(1.0451)	(1.3292)	(3.1752)	(1.3658)	(2.8896)
Obs	600	400	600	400	600	400
Clusters	24	16	24	16	24	16

Notes: Clustered standard errors at group level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients correspond to the specification model (6) using data from treatments T1–T3, with T1 as the baseline (i.e., T1 serves as the reference group for comparisons). Columns (1), (3), and (5) present random-effects GLS regressions on fiscal policy, with standard errors clustered at the group level. Columns (2), (4), and (6) restrict the sample to treatments T1 and T3, including two additional control variables: the gender of the winning candidate and the group's gender composition.

positive and statistically significant effect ($p < 0.05$) in Column (2). To contextualize this effect, consider the baseline corruption level represented by the constant term in Column (2). This baseline corruption level of 20.504 ECUs corresponds to a group with no females. For a group with one female, the expected corruption level can be calculated as: Corruption = $20.504 + 0.401 \times 1 = 20.905$ ECUs. Similarly, for a group with, for instance, three females, the expected corruption level increases further to Corruption = $20.504 + 0.401 \times 3 = 21.707$ ECUs. Thus, each additional female in the group is associated with an increase of approximately 0.401 ECUs (around 2%) in the corruption level chosen by the winning candidate.

This finding suggests that gender dynamics within groups play a significant role in influencing fiscal policy decisions related to corruption. Specifically, a higher representation of females in the group is associated with higher corruption levels, contrary to some theoretical expectations that increased female representation might lead to lower corruption due to different ethical perspectives.

Taxes. When analysing Taxes as the dependent variable, the coefficient for Round is strongly positive and highly statistically significant ($p < 0.01$) in both models. This means that each additional round leads to a significant increase in the tax level chosen by the winning candidate.

Treatment T2 (Automated Voters) has a positive and statistically significant effect on Taxes (3.525, $p = 0.041$), indicating that automated voters significantly increase the tax level compared to the baseline Treatment T1. Treatment T3 (Full Transparency) also shows a positive and statistically significant effect (2.710, $p = 0.048$ in Column (3) and 3.211, $p = 0.006$ in Column (4)), suggesting that full transparency significantly increases the tax level compared to T1.

The constant term in both models is highly significant ($p < 0.01$), representing the baseline tax rate in Treatment T1. When compared with the Nash Equilibrium prediction of 52.52 ECUs, the baseline tax level in Column (4) (52.148) aligns closely with the equilibrium prediction. However, Column (3) (58.137) is 11% higher. Treatments T2 and T3 further increase the tax levels beyond the equilibrium, indicating that automated voters and transparency encourage slightly higher tax proposals.

The control variable Group's Gender Composition has a positive and statistically significant effect (1.290, $p = 0.018$) in Column (4), indicating that a higher female presence within the group significantly increases the tax level chosen by the winning candidate. By contrast, the Gender of Winning Candidate shows a marginally significant negative effect (-1.676, $p = 0.091$) in Column (4), suggesting that female winning candidates may select slight lower tax rates than their male counterparts, though this effect is not strongly significant.

Public good. For the Public Good dependent variable, the coefficient for Round is strongly positive and highly statistically significant in both models (0.367, $p < 0.01$ and 0.445, $p < 0.01$). This indicates that each additional round leads to a significant increase in the public good level chosen by the winning candidate.

Treatment T2 (Automated Voters) shows a positive and marginally significant effect (4.555, $p = 0.074$) in Column (5), suggesting that automated voters tend to increase the public good level, although the effect is only marginally significant. Treatment T3 (Full Transparency) has a positive and marginally significant effect (2.327, $p = 0.073$) in Column (6), indicating a similar tendency under full fiscal transparency.

None of the control gender variables have statistically significant effects on Public Good. Finally, the constant term in both models (36.145 without controls and 31.932 with controls) are highly significant ($p < 0.01$), and they represent the baseline public good level in Treatment T1. These observed values are below the Nash Equilibrium

prediction of 48 ECUs. Treatments T2 and T3 show tendencies to increase public good levels, suggesting that automated voters and full transparency encourage higher public good allocations. However, these effects are only marginally significant and do not fully bridge the gap to the equilibrium.

Having established that Treatments T1, T2, and T3 converge to similarly high levels of corruption, we now turn our attention to Treatment 4. Unlike T1–T3, T4 introduces a *disproportional* power-sharing rule, fundamentally altering the theoretical equilibrium. Under this treatment, the predicted corruption level drops significantly to approximately 1.38, compared to 4.04 in the other treatments. This stark reduction highlights a major shift in the strategic dynamics of the game: the disproportional rule intensifies the competition between candidates, as securing a majority of votes now yields a disproportionately large share of policy-making power. This creates stronger incentives for candidates to reduce corruption and prioritize voter accountability, potentially fostering a more equitable policy environment.

To empirically assess the differences between the baseline Treatment T1 and the alternative T4, we estimate the following regression model:

$$\text{Fiscal Policy}_{g,t} = \alpha + u_g + \beta \cdot \text{T4}_g + \gamma_1 \cdot \text{Round}_t + \gamma_2 \cdot \text{Controls}_{g,t} + \varepsilon_{g,t}, \tag{7}$$

where the variable T4_g is a dummy indicating Treatment 4, the term $\text{Controls}_{g,t}$ captures relevant demographic characteristics of the participants, i.e., the gender of the winning candidate and the gender composition of the group, and the rest is defined as before.[14] Group fixed effects are excluded to prevent collinearity arising from single-treatment groups, and standard errors are clustered at the group level to account for within-group correlation and heteroscedasticity.

The dataset consists of 400 observations, offering robust statistical power to identify and evaluate the effects of the disproportional power-sharing rule in Treatment 4. The key findings from this analysis are presented in Table 5, accompanied by a detailed discussion and interpretation of the results.

Corruption. In the regression models where Corruption is the dependent variable, the coefficient for Round is strongly negative and highly statistically significant (-0.175, $p < 0.01$ and -0.176, $p < 0.01$ in Columns (1) and (2), respectively). This indicates that with each additional round, the corruption level chosen by the winning candidate decreases by approximately 0.175 ECUs, holding all other factors con-

[14] In T1 and T4, there are 231 females (57%) among 400 winning candidates. 201 (50.3%) groups have 4 or less females, and 199 (49.7%) groups have more than 4 females in the group.

Table 5: Winning Candidate's Fiscal Policy (Treatments T1 & T4)

	Corruption$_{gt}$		Taxes$_{gt}$		Public Good$_{gt}$	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Round$_t$	-0.175***	-0.176***	0.195*	0.195*	0.370***	0.371***
	(0.0599)	(0.0603)	(0.1067)	(0.1072)	(0.0930)	(0.0948)
T4$_g$	-6.000***	-5.922***	2.535	2.497	8.535***	8.419***
	(0.8886)	(0.9459)	(1.9857)	(2.0062)	(1.9387)	(2.0264)
Winner Gender$_{gt}$		1.245		-0.109		-1.389
		(0.9852)		(0.6938)		(1.2748)
Group Gender$_{gt}$		0.038		-0.127		-0.159
		(0.2431)		(0.5423)		(0.4577)
Constant	23.854***	22.944***	59.953***	60.591***	36.099***	37.641***
	(0.7661)	(1.3609)	(1.7517)	(3.4011)	(1.5776)	(2.9934)
Obs	400	400	400	400	400	400
Clusters	16	16	16	16	16	16

Notes: Clustered standard errors at group level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients correspond to the specification model (7), with data from treatments T1 and T4 (T1 as baseline). Columns **(1)**, **(3)** and **(5)** report the random-effects GLS regressions on fiscal policy, clustering standard errors by group. Columns **(2)**, **(4)** and **(6)** adds additional controls, namely, gender of the winning candidate, and group gender composition.

stant. This negative relationship suggests that as the number of rounds progresses, candidates may become less inclined to engage in corrupt practices, possibly due to increased scrutiny or strategic considerations over time.

For Treatment T4 (Disproportional Power-Sharing), the coefficient is strongly negative and highly significant (-6.000, $p < 0.01$ and -5.922, $p < 0.01$ in Columns (1) and (2), respectively). This indicates that Treatment T4 reduces the corruption level chosen by the winning candidate by approximately 6 ECUs compared to the baseline Treatment T1. This substantial reduction underscores the effectiveness of disproportional power-sharing in mitigating corruption, likely by increasing accountability and aligning power incentives more closely with ethical governance.

The control variable group gender composition does not have a statistically significant effect on Corruption in Column (2). This suggests that, unlike in Treatments T1–T3, having more females within the group does not significantly influence the corruption level chosen by the winning candidate in the context of Treatments T1 and T4. Similarly, the gender of winning candidate remains not significant, indicating that it does not play a significant role in determining corruption under these treatment conditions.

The constant term in both models is highly significant ($p < 0.01$), representing the estimated baseline corruption in Treatment T1. Compared to the Nash Equilibrium level of approximately 4.04 ECUs, the empirical findings reveal that the baseline corruption levels—around 23.854 (without controls) and 22.944 (with controls)—are substantially higher than the equilibrium prediction. This stark divergence indicates that candidates in the experiment consistently choose corruption levels far exceeding theoretical expectations, even after accounting for the effects of Treatment T4. While disproportional power-sharing reduces corruption by approximately 6 ECUs, resulting in levels of 17.854 and 16.944 ECUs, these remain significantly above the theoretical benchmark, underscoring the persistence of high corruption in our experiment.

Taxes. When analysing Taxes as the dependent variable, the coefficient for Round is positive and marginally significant (0.195, $p < 0.10$ in both Columns (3) and (4)). This means that each additional round leads to a slight but marginally significant increase in the tax level chosen by the winning candidate, suggesting a trend towards higher taxation as the game progresses.

Treatment T4 (Disproportional Power-Sharing) has a positive but not statistically significant effect on Taxes. This indicates that unlike Treatments T1–T3, where certain treatments significantly influenced tax levels, disproportional power-sharing in the context of Treatments T1 and T4 does not have a meaningful impact on taxation. Similar to Corruption, the control gender variables does not significantly influence tax levels in the context of Treatments T1 and T4.

The constant term in both models (59.953 and 60.591) is highly significant ($p < 0.01$), representing the estimated baseline tax rate in Treatment T1. When compared with the Nash Equilibrium prediction of approximately 52.52 ECUs, the baseline tax levels in both columns exceed the equilibrium value. Treatment T4 does not significantly alter these tax levels, indicating that disproportional power-sharing does not further drive tax increases beyond the already elevated baseline observed in T1.

Public good. For the Public Good dependent variable, the coefficient for Round is strongly positive and highly statistically significant (0.370, $p < 0.01$ and 0.371, $p < 0.01$ in Columns (5) and (6), respectively). This indicates that each additional round leads to a significant increase in the public good level chosen by the winning candidate, suggesting a consistent trend towards greater investment in public goods as the game progresses.

Treatment T4 (Disproportional Power-Sharing) exhibits a strong positive and highly significant effect (8.535, $p < 0.01$ and 8.419, $p < 0.01$ in Columns (5) and (6),

respectively). This substantial increase of approximately 8.5 ECUs in the public good level compared to the baseline Treatment T1 underscores the effectiveness of disproportional power-sharing in promoting greater public good provision. The significant positive impact indicates that when policy-making power is allocated more disproportionally to candidates with over 50% vote share, they are more likely to prioritize investments in public goods, likely due to enhanced accountability and the strategic benefits of supporting public welfare in gaining sustained electoral support.

Finally, the constant term in both models (36.099 without controls and 37.641 with controls) is highly significant ($p < 0.01$). These observed baseline values are below the Nash Equilibrium (48.48 in T1 and 59.22 in T4). Although Treatment T4 significantly increases public good levels by approximately 8.5 ECUs, the resulting levels (44.634 and 46.060 ECUs) still remain below the theoretical prediction. This indicates that while disproportional power-sharing promotes greater investment in public goods, other factors may be influencing the remaining gap towards equilibrium levels.

Overall, while Treatment T4 (Disproportional Power-Sharing) effectively reduces corruption and enhances public good provision, it does not significantly influence tax levels nor does it align corruption and public good levels fully with Nash Equilibrium predictions. The lack of significant gender-related effects in Treatments T1 and T4 further suggests that gender dynamics may interact with institutional settings in complex ways. These findings highlight the importance of institutional power structures in shaping fiscal policy decisions while also indicating the need for additional strategies to address persistent corruption and achieve higher public good provision.

To complement the take-away from Table 5, the regression results presented in Table 6 examine whether the differences in fiscal policy decisions between Treatments T1 and T4 evolve over time. This is achieved through the interaction term ($T4_g \times$ Round), which captures whether the trajectory of corruption, taxes, and public good provision over time varies by treatment. The main findings are as follows:

Corruption. The interaction term for corruption is negative and statistically significant (-0.245, $p < 0.05$), indicating that the reduction in corruption under Treatment 4 intensifies as the experiment progresses. This result implies that the gap in corruption between T1 and T4 widens over time, with participants in T4 exhibiting a steeper decline in corruption as they adjust to the treatment's institutional framework.[15]

[15]It is important to remind that our experiment included fewer rounds than other studies in the literature. This is due to the use of human voters, rather than computer algorithms, which required additional time to complete each round. Thus, subjects may not have had sufficient time to fully converge to the equilibrium values, potentially affecting the observed learning process.

Taxes. The interaction term for taxes is also negative and marginally significant (-0.352, $p < 0.10$). While Treatment 4 initially leads to higher taxes (main effect: 7.116, $p < 0.10$), the increase diminishes over time, suggesting that the gap in tax levels between T1 and T4 narrows as the session progresses. This could indicate that participants in T4 adjust their tax strategies or voters respond differently over time.

Public good. The interaction term for public good provision is negative but not statistically significant (-0.108, $p > 0.10$), suggesting that the observed increase in public good provision under T4 (9.936, $p < 0.01$) is consistent across all rounds. This lack of significant interaction indicates that the difference in public good provision between T1 and T4 remains stable throughout the session.

Table 6: Interaction of Treatment 4 and Round Effects on Fiscal Policy

	Corruption_{gt}	Taxes_{gt}	Public Good_{gt}
Round_t	-0.053	0.371***	0.424***
	(0.0883)	(0.1109)	(0.1132)
T4_g	-2.820*	7.116*	9.936***
	(1.6494)	(3.7686)	(3.0087)
$\text{T4}_g \times \text{Round}_t$	-0.245**	-0.352*	-0.108
	(0.1019)	(0.1932)	(0.1842)
Constant	22.264***	57.663***	35.399***
	(1.0431)	(1.9719)	(1.8545)
Obs	400	400	400
Clusters	16	16	16

Notes: Clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients correspond to random-effects GLS regressions on fiscal policy, clustering standard errors by group, with data from treatments T1 and T4 (T1 as baseline).

6.2 Payoff differences and voting behavior

Continuing with the data analysis, we now examine whether voters align their choices with direct monetary incentives, as predicted by the theoretical model. Specifically, we focus on ΔUtility, a key predictor in the regression specifications, to assess whether voters respond to payoff differences or whether other factors drive their decisions.

ΔUtility represents the difference in monetary (ECUs) payoffs between choosing candidate A and candidate B. This measure captures the net monetary benefit a voter

would receive by selecting candidate A, accounting for differences in the fiscal policies proposed by the two candidates. By incorporating ΔUtility into our analysis, we aim to test whether voters prioritize material benefits over ideological or non-monetary considerations in their electoral choices.

The logistic regression specification we run can be expressed as follows:

$$\begin{aligned}\text{logit}\Big(P(\text{Vote}_{igt}(A)=1)\Big) \;=\; & \alpha + \beta_1 \cdot \Delta\text{Utility}_{igt} + \beta_2 \cdot \text{Partisan Bias}_{it} \\ & + \gamma_1 \cdot \text{T2}_g + \gamma_2 \cdot \text{T3}_g + \gamma_3 \cdot \text{T4}_g \\ & + \delta_1 \cdot (\text{T2}_g \times \Delta\text{Utility}_{igt}) + \delta_2 \cdot (\text{T3}_g \times \Delta\text{Utility}_{igt}) \\ & + \delta_3 \cdot (\text{T4}_g \times \Delta\text{Utility}_{igt}) + \phi_1 \cdot (\text{Partisan Bias}_{it} \times \text{T2}_g) \\ & + \phi_2 \cdot (\text{Partisan Bias}_{it} \times \text{T3}_g) + \phi_3 \cdot (\text{Partisan Bias}_{it} \times \text{T4}_g) \\ & + \theta_1 \cdot \text{Voter Gender}_i + \theta_2 \cdot \text{Round FE}_t + \varepsilon_{igt},\end{aligned} \tag{8}$$

where:

- $\text{Vote}_{igt}(A)$ is a binary outcome variable indicating whether voter i chooses candidate A (1) or candidate B (0) in group g at time t,
- $\Delta\text{Utility}_{igt}$ denotes voter i's utility difference (expressed in ECUs) from the fiscal policies of candidate A and candidate B in group g at time t,
- T2_g, T3_g, T4_g are binary indicators for Treatments T2, T3 and T4, respectively,
- $\text{Partisan Bias}_{it}$ is a continuous variable measuring the voter i's predisposition towards a specific candidate in period t,
- Voter Gender_i is a binary variable indicating the gender of voter i, used to control for potential gender-specific voting patterns,
- Round FE_t denotes fixed effects for each experimental round to account for unobserved time-specific heterogeneity across election rounds, and
- ε_{igt} is an error term capturing unobserved influences on voting decisions.

The coefficient β_1 measures the overall responsiveness of voters to payoff differences. A significant β_1 would indicate that voters favour candidates offering higher utility payoffs, aligning with rational decision-making models. Conversely, insignificant values suggest that other factors, such as partisan bias, may overshadow monetary considerations. The coefficient β_2 captures the direct effect of partisan bias on

voter preferences. In the special case where $\beta_2 = -\beta_1$, the model reduces to a simpler form, $\beta_1 \cdot (\Delta\text{Utility}_{igt} - \text{Partisan Bias}_{it})$, which aggregates material and ideological considerations into a single term.

This special case captures the combined effects of utility differences and ideological biases on voters' decisions, providing a unified measure of material and partisan considerations. Columns (1) and (4) of Table 7 analyse this aggregated effect for the two groups of treatments: T1–T3 and T1 & T4, respectively. These results establish a baseline for understanding the relative importance of these variables. The remaining columns, (2)–(3) and (5)–(6), disaggregate the effects by examining cases where $\beta_2 \neq -\beta_1$, allowing for a more nuanced comparison of the distinct roles played by material and partisan factors across experimental conditions.

Control variables, such as voter gender and round fixed effects, are included in all regressions. They are intended to account for individual and contextual factors that might influence voting behavior. Additionally, interaction terms test whether voter responsiveness to payoff differences varies across treatments. For instance, a significant δ_1 indicates that voters in T2 (automated voters) respond differently to ΔUtility compared to T1. Similarly, the interaction terms for partisan bias assess whether institutional treatments influence the weight voters place on ideological preferences.

Table 7 presents estimates from logistic regressions that assess how the combined effect of monetary incentives and partisan bias, as well as their separate components and interactions with treatment conditions, influence the probability of a voter choosing candidate A. The table is organized in two sets: the first three columns report results from Treatments T1–T3 (with T1 as the baseline), and the latter three columns report results from a pooled sample of Treatments T1 and T4. In both sets the key aggregated predictor is "$\Delta\text{Utility}_{igt}$ – Partisan Bias_{it}," while separate coefficients for $\Delta\text{Utility}_{igt}$ and Partisan Bias_{it} allow to examine their individual contributions.

In the T1–T3 sample, the aggregated coefficient for ΔUtility – Partisan Bias is estimated at approximately 6.43 (Column (1)), and a similar magnitude (6.32) is found in the T1 & T4 specification (Column (4)). This indicates that, holding all else constant, a one-unit increase in the net monetary benefit for candidate A (after accounting for partisan bias) is associated with an increase of about 6.4 log-odds in the probability of voting for A.

When the components are estimated separately (Columns (2) and (3) for T1–T3), the coefficient on ΔUtility is positive and significant (around 3.89 and 3.79, respectively), whereas the coefficient for Partisan Bias is strongly negative (approximately −12.00 and −10.88). This contrast demonstrates that monetary incentives tend to increase support for candidate A, but partisan bias exerts a large adverse effect (see Fig-

Table 7: Payoff Differences and Voting for A: Logistic Regression

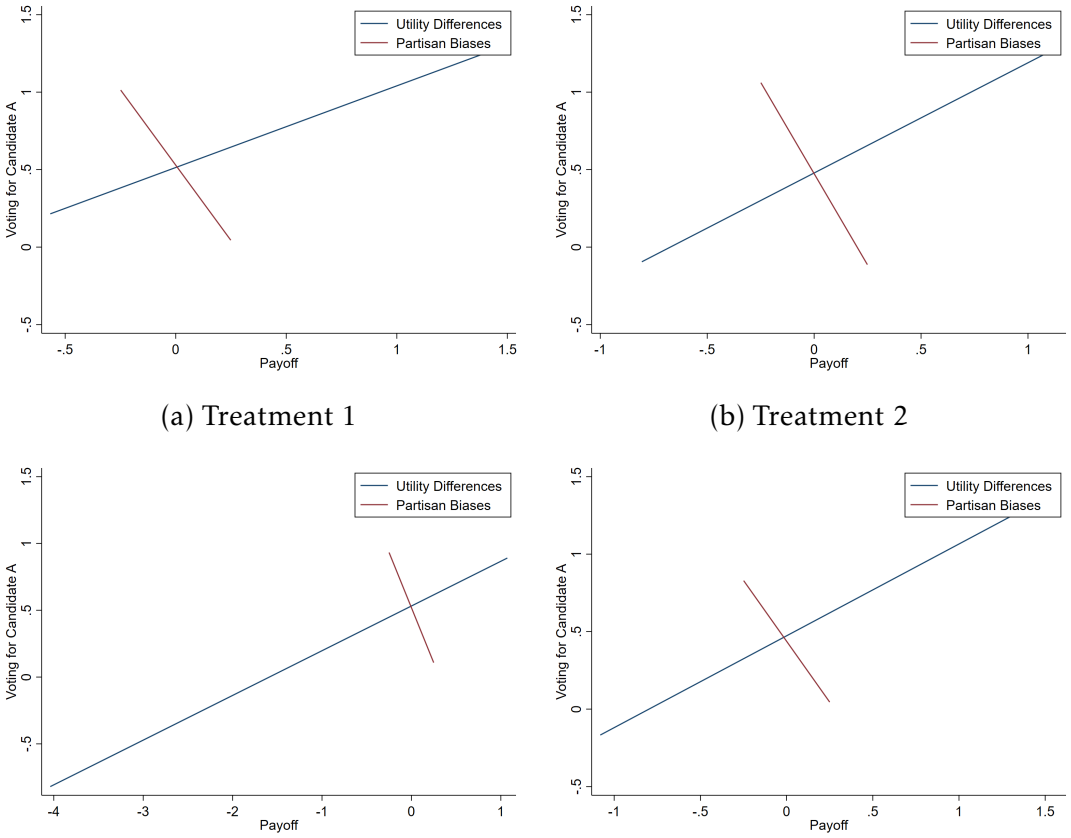
$\text{Vote}_{igt}(A)$ (Dep. Var.)	**Treatments T1–T3**			**Treatments T1 & T4**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$\Delta\text{Utility}_{igt} - \text{Partisan Bias}_{it}$	6.426***			6.319***		
	(1.2758)			(0.5126)		
$T2_g$	0.002	-0.140	-0.190			
	(0.1461)	(0.1623)	(0.2093)			
$T3_g$	0.250*	0.211	0.158			
	(0.1419)	(0.1508)	(0.1520)			
$T4_g$				-0.045	-0.192	-0.182
				(0.1350)	(0.1529)	(0.1573)
$\Delta\text{Utility}_{igt}$		3.891***	3.787***		3.941***	3.852***
		(1.0144)	(1.0610)		(0.4815)	(1.0835)
$\text{Partisan Bias}_{it}$		-12.000***	-10.880***		-9.739***	-10.938***
		(1.1525)	(0.6298)		(0.7090)	(0.6225)
$T2_g \times \Delta\text{Utility}_{igt}$			6.054**			
			(2.7038)			
$T3_g \times \Delta\text{Utility}_{igt}$			-1.333			
			(1.3528)			
$T4_g \times \Delta\text{Utility}_{igt}$						0.052
						(1.1634)
$\text{Partisan Bias}_{it} \times T2_g$			-11.824***			
			(4.0603)			
$\text{Partisan Bias}_{it} \times T3_g$			2.126			
			(1.3083)			
$\text{Partisan Bias}_{it} \times T4_g$						2.283*
						(1.1877)
Voter Gender_i	-0.051	-0.021	-0.031	-0.163	-0.154	-0.146
	(0.0866)	(0.0964)	(0.0927)	(0.1037)	(0.1102)	(0.1088)
Round FE_t	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.321	0.391	0.511	0.820	0.908	0.906
	(0.3490)	(0.3564)	(0.3976)	(0.5834)	(0.5642)	(0.5953)
Obs	3000	3000	3000	2000	2000	2000
Clusters	24	24	24	16	16	16

Notes: Clustered standard errors at group level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ures 2a–2d for graphical visualization). The aggregated effect of these two channels, roughly the sum of the coefficients on ΔUtility and Partisan Bias, is therefore positive when ΔUtility outweighs the negative influence of the ideological bias.

Regarding treatment effects, the dummy for Treatment T2 does not show significant differences from T1 (with coefficients near zero or only marginally negative), whereas T3 appears to exert a mild positive effect (around 0.250 in Column (1), significant at the 10% level). In contrast, in the T1 & T4 specification, the T4 dummy is negative (ranging from −0.045 to −0.192), though these estimates are not significant across specifications. Thus, for these models, the primary impact on voters' choices

Figure 2: The Impact of Utility Difference and Partisan Bias on Voting for *A*.



(a) Treatment 1

(b) Treatment 2

(c) Treatment 3

(d) Treatment 4

seems to stem from the monetary and partisan components rather than directly from the treatment dummies.

The interaction terms shed further light on treatment-specific moderation. In the T1–T3 models (Column (3)), the coefficient on $T2_g \times \Delta\text{Utility}$ is approximately 6.05 (significant at the 5% level), indicating that in T2, relative to T1, the effect of a one-unit increase in ΔUtility on voting for *A* is strongly amplified. Likewise, the interaction of Partisan Bias$\times T2_g$ is strongly negative (−11.82), and highly significant at the 1% level, suggesting that T2 also amplifies the adverse effect of partisan bias. One possible interpretation is that voters in the automated treatment pay more attention to both monetary gains *and* ideological preferences, potentially because automation changes the perceived trade-off between material benefits and partisan alignment. Meanwhile, the interaction of $T3_g \times \Delta\text{Utility}$ is small and not statistically significant, indicating that T3 does not modify the relationship between monetary incentives and voting.

In the pooled T1 & T4 models, the interaction between Partisan Bias and $T4_g$ is moderately positive (around 2.283, significant at the 10% level), implying that the

negative influence of partisan bias is mitigated in T4 relative to T1. This mitigation could reflect how disproportional power-sharing affect voter perceptions, making them less rigid in adhering to partisan preferences when material benefits are at stake. Importantly, neither the dummy for T4 nor its interaction with ΔUtility reaches statistical significance, suggesting that any direct effect of T4 on the likelihood of voting for A is overshadowed by how it modifies the weight of partisan bias.

Finally, the control variable Voter Gender consistently has a small and statistically insignificant effect across all specifications, implying that gender does not play a major role in the decision to vote for candidate A. The constant terms, which reflect the baseline log-odds of voting for A (when all predictors are zero), vary across models but mainly serve as a reference point for interpreting the other coefficients.

Overall, the results suggest that while both fiscal incentives (reflected in ΔUtility) and partisan bias are important determinants of voting behavior, partisan bias has a particularly strong negative effect. Moreover, the way treatment conditions interact with these variables indicates that institutional settings (e.g., automation or changes in power-sharing rules) can alter the relative salience of monetary incentives versus political leanings. Treatment T2 appears to amplify the weight of both factors, whereas T4 somewhat reduces the negative impact of partisan bias, highlighting how institutional details can shape the interplay between economic and partisan preferences.

6.3 Policy differences and voting behavior

In the analysis of voting behavior, we began examining the effects of payoff differences before turning, in this section, to how each policy dimension—taxes, corruption, and public goods—individually influences voter choices. This approach reflects the interconnected nature of the fiscal variables, ensuring a logical progression and a systematic framework for understanding their aggregate and individual impacts.

In Section 6.2, the variable ΔUtility was introduced as a composite measure aggregating the differences in fiscal policy components, offering a holistic view of the net monetary incentives shaping voter behavior. By focusing initially on ΔUtility, we captured the overall impact of combined policy differences between candidates A and B on voter decisions. This aggregated perspective aligns with the theoretical model's emphasis on the combined effects of multiple policy dimensions, providing an understanding of how integrated fiscal incentives drive electoral outcomes.

In contrast, this section adopts a more disaggregated approach, isolating specific components of fiscal policy. This approach accounts for the intrinsic links between taxes, corruption, and public goods through the budget constraint. Due to these interdependencies, including all three components in a single regression would introduce

multicollinearity, compromising the reliability of coefficient estimates. To address this, our analysis examines pairs of policy components individually, mitigating multicollinearity and providing clearer insights into how each fiscal dimension independently influences voter choices.

The logistic regression model we estimate is as follows:

$$\begin{aligned}\text{logit}\Big(P(\text{Vote}_{igt}(A)=1)\Big) \;=\; & \alpha + \beta_1 \cdot \Delta\text{Corruption}_{gt} + \beta_2 \cdot \Delta\text{Taxes}_{gt} + \beta_3 \cdot \Delta\text{Public Good}_{gt} \\ & + \gamma_1 \cdot (\text{T2}_g \times \Delta\text{Corruption}_{gt}) + \gamma_2 \cdot (\text{T3}_g \times \Delta\text{Corruption}_{gt}) \\ & + \delta_1 \cdot (\text{T2}_g \times \Delta\text{Taxes}_{gt}) + \delta_2 \cdot (\text{T3}_g \times \Delta\text{Taxes}_{gt}) \\ & + \phi_1 \cdot (\text{T2}_g \times \Delta\text{Public Good}_{gt}) + \phi_2 \cdot (\text{T3}_g \times \Delta\text{Public Good}_{gt}) \\ & + \theta_1 \cdot \text{Partisan Bias}_{it} + \theta_2 \cdot \text{Initial Income}_{igt} + \theta_3 \cdot \text{Voter Gender}_i \\ & + \theta_4 \cdot \text{Round FE}_t + \varepsilon_{igt}\end{aligned} \tag{9}$$

where:

- $\Delta\text{Corruption}_{gt}$ is the difference in corruption associated with candidates' policies in group g at election time t,
- ΔTaxes_{gt} denotes the difference in the tax level proposed by candidates in group g at time t
- $\Delta\text{Public Good}_{gt}$ is the difference in the provision of public goods resulting from candidate proposals in group g at election time t,
- $\text{Initial Income}_{igt}$ is a continuous variable representing each voter i's initial income level in group g at period t,

and the rest of the terms in (9) are defined as in equation (8).

Table 8 presents the regression results. The analysis is pooled across treatments, using Treatment T1 as the reference category. Columns (1) through (9) correspond to distinct model specifications, incrementally exploring pairs of fiscal variables and incorporating interaction terms and control variables as described above.

The regressions deliver several interesting insights into voter behavior. First, the main effects reveal robust patterns in how voters respond to fiscal policy differences. Increases in corruption proposed by candidate A relative to candidate B (ΔCorruption) significantly decrease the likelihood of voters choosing A, with this effect becoming stronger across later models (e.g., from -0.042 in Column 2 to -0.070 in Column 9, $p < 0.01$). Similarly, higher taxes proposed by candidate A (ΔTaxes) also reduce voter support, as shown by significant negative coefficients ($p < 0.01$) in early models like

Table 8: Fiscal Policy Effects on Voting for A: Logistic Regression (T1–T3 Pooled).

$\text{Vote}_{igt}(A)$ (Dep. Var.)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
$\Delta\text{Corruption}_{gt}$		-0.042***	-0.045***		-0.041***	-0.043***		-0.066***	-0.070***
		(0.0076)	(0.0078)		(0.0074)	(0.0089)		(0.0118)	(0.0136)
ΔTaxes_{gt}	-0.042***		0.002	-0.041***		0.002	-0.066***		0.004
	(0.0076)		(0.0022)	(0.0074)		(0.0048)	(0.0118)		(0.0066)
$\Delta\text{Public Good}_{gt}$	0.045***	0.002		0.043***	0.002		0.070***	0.004	
	(0.0078)	(0.0022)		(0.0089)	(0.0048)		(0.0136)	(0.0066)	
$T2_g \times \Delta\text{Corruption}_{gt}$					-0.022**	-0.015		-0.034**	-0.021
					(0.0108)	(0.0119)		(0.0163)	(0.0163)
$T3_g \times \Delta\text{Corruption}_{gt}$					0.009	0.006		0.016	0.013
					(0.0124)	(0.0141)		(0.0193)	(0.0219)
$T2_g \times \Delta\text{Taxes}_{gt}$				-0.022**		-0.007	-0.034**		-0.013
				(0.0108)		(0.0070)	(0.0163)		(0.0096)
$T3_g \times \Delta\text{Taxes}_{gt}$				0.009		0.004	0.016		0.004
				(0.0124)		(0.0053)	(0.0193)		(0.0074)
$T2_g \times \Delta\text{Pub.Good}_{gt}$				0.015	-0.007		0.021	-0.013	
				(0.0119)	(0.0070)		(0.0163)	(0.0096)	
$T3_g \times \Delta\text{Pub.Good}_{gt}$				-0.006	0.004		-0.013	0.004	
				(0.0141)	(0.0053)		(0.0219)	(0.0074)	
$\text{Partisan Bias}_{it}$							-12.237***	-12.237***	-12.237***
							(1.0847)	(1.0847)	(1.0847)
$\text{Initial Income}_{igt}$							-0.011	-0.011	-0.011
							(0.0153)	(0.0153)	(0.0153)
Voter Gender_i							0.030	0.030	0.030
							(0.1029)	(0.1029)	(0.1029)
Round FE_t	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.294	0.294	0.294	0.412**	0.412**	0.412**	0.894*	0.894*	0.894*
	(0.2033)	(0.2033)	(0.2033)	(0.2055)	(0.2055)	(0.2055)	(0.5260)	(0.5260)	(0.5260)
Obs	3000	3000	3000	3000	3000	3000	3000	3000	3000
Clusters	24	24	24	24	24	24	24	24	24

Notes: Clustered standard errors at group level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients correspond to the specification model (9), using data from Treatments T1–T3, with T1 as the baseline. Columns **(1)**–**(9)** present logistic regression models with varying combinations of interaction terms and control variables. Controls include Partisan Bias, Individual Income, Voter Gender, and Round Fixed Effects.

Columns 1, 4, and 7. However, this effect diminishes in significance when interaction terms and additional controls are introduced. In contrast, increases in public good provision (ΔPublic Good) consistently lead to a higher probability of voters choosing candidate A. This effect strengthens across models, with coefficients rising from 0.045 in Column 1 to 0.070 in Column 7 ($p < 0.01$).

Interaction effects further highlight treatment-specific dynamics in voter behavior. For automated voters in T2, the penalty for corruption is amplified, as evidenced by the significant and negative interaction term ($T2_g \times \Delta\text{Corruption}$) in Columns 5 and 8 ($p < 0.05$). A similar pattern emerges for taxes, where the interaction term $T2_g \times \Delta\text{Taxes}$ shows that voters in T2 impose stricter penalties on higher taxes compared to T1 ($p < 0.05$ in Columns 4 and 7). However, other interaction terms, such as $T3_g \times \Delta\text{Corruption}$ and $T2_g \times \Delta\text{Public Good}$, are not statistically significant, suggesting no systematic differences in voter responses to these policies across treatments.

Among the control variables, partisan bias emerges as a strongly negative and significant predictor ($p < 0.01$), indicating that voters with higher bias against A are substantially less likely to support this candidate. In contrast, other controls, such as initial income and voter gender, show no significant influence on voting decisions, pointing to the limited role of these factors in our experimental setting.

Treatment-specific effects also underscore the distinct roles of institutional settings. Automated voters in T2 show stricter penalties for both corruption and taxes, amplifying the negative effects of ΔCorruption and ΔTaxes relative to T1. On the other hand, full transparency in T3 does not significantly alter voter sensitivities to fiscal policy differences, as evidenced by the non-significant interaction terms involving $T3_g$.

In conclusion, the main takeaway from Table 8 is that voters penalize candidates proposing higher corruption and taxes while rewarding those offering greater public good provision. This pattern is intensified under T2 due to automated voters but remains largely unchanged under T3. Lastly, the persistent significance of partisan bias highlights the critical role of individual ideology in shaping electoral choices, regardless of fiscal policies or institutional treatments.

Next, we examine voter behavior across Treatments T1 and T4, as displayed in Table 9. The main effects shows that increases in corruption by candidate A relative to candidate B (ΔCorruption) significantly reduce the probability of voters choosing candidate A in all models ($p < 0.01$). For instance, the coefficient is −0.050 in Column (2) and −0.066 in Column (9), indicating that voters are less likely to support candidates associated with higher corruption. This pattern is consistent with the findings for T1–T3 but appears slightly stronger in later models for T1 and T4, reinforcing the idea that voters consistently punish corruption.

Similarly, higher taxes proposed by A (ΔTaxes) also decrease voter support, with significant and negative coefficients ($p < 0.01$) in early models such as Columns (1), (4), and (7). However, this effect becomes non-significant in later models, suggesting that the influence of taxes on voter decisions diminishes when additional controls and interaction terms are included. This pattern mirrors the results observed in T1–T3, though the absolute effects are slightly smaller in the T1 and T4 pooled results. On the other hand, increases in public good provision (ΔPublic Good) significantly boost voter support for candidate A across all relevant models ($p < 0.01$). The effect size is consistent with or slightly larger than the corresponding effects in T1–T3, as seen in coefficients such as 0.054 in Column (1) and 0.066 in Column (7). This finding underscores voters' consistent preference for higher public good provision.

The interaction effects between T4 and fiscal policy variables reveal no statisti-

Table 9: Fiscal Policy Effects on Voting for A: Logistic Regression (T1 & T4 Pooled).

$Vote_{igt}(A)$ (Dep. Var.)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
$\Delta Corruption_{gt}$		-0.050***	-0.054***		-0.044***	-0.047***		-0.062***	-0.066***
		(0.0057)	(0.0059)		(0.0073)	(0.0087)		(0.0100)	(0.0114)
$\Delta Taxes_{gt}$	-0.050***		0.004	-0.044***		0.003	-0.062***		0.004
	(0.0057)		(0.0037)	(0.0073)		(0.0050)	(0.0100)		(0.0058)
$\Delta Public\ Good_{gt}$	0.054***	0.004		0.047***	0.003		0.066***	0.004	
	(0.0059)	(0.0037)		(0.0087)	(0.0050)		(0.0114)	(0.0058)	
$T4_g \times \Delta Corruption_{gt}$					-0.010	-0.013		-0.011	-0.015
					(0.0125)	(0.0121)		(0.0157)	(0.0153)
$T4_g \times \Delta Taxes_{gt}$				-0.010		0.002	-0.011		0.004
				(0.0125)		(0.0067)	(0.0157)		(0.0077)
$T4_g \times \Delta Pub.\ Good_{gt}$				0.013	0.002		0.015	0.004	
				(0.0121)	(0.0067)		(0.0153)	(0.0077)	
$Partisan\ Bias_{it}$							-9.977***	-9.977***	-9.977***
							(0.6535)	(0.6535)	(0.6535)
$Initial\ Income_{igt}$							0.009	0.009	0.009
							(0.0162)	(0.0162)	(0.0162)
$Voter\ Gender_i$							-0.122	-0.122	-0.122
							(0.1213)	(0.1213)	(0.1213)
Round FE_t	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.239	0.239	0.239	0.237	0.237	0.237	0.363	0.363	0.363
	(0.3008)	(0.3008)	(0.3008)	(0.2894)	(0.2894)	(0.2894)	(0.6363)	(0.6363)	(0.6363)
Obs	2000	2000	2000	2000	2000	2000	2000	2000	2000
Clusters	16	16	16	16	16	16	16	16	16

Notes: Clustered standard errors at group level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients correspond to the specification model (9), using data from Treatments T1 and T4, with T1 as the baseline. Columns **(1)–(9)** present logistic regression models with varying combinations of interaction terms and control variables. Controls include Partisan Bias, Individual Income, Voter Gender, and Round Fixed Effects.

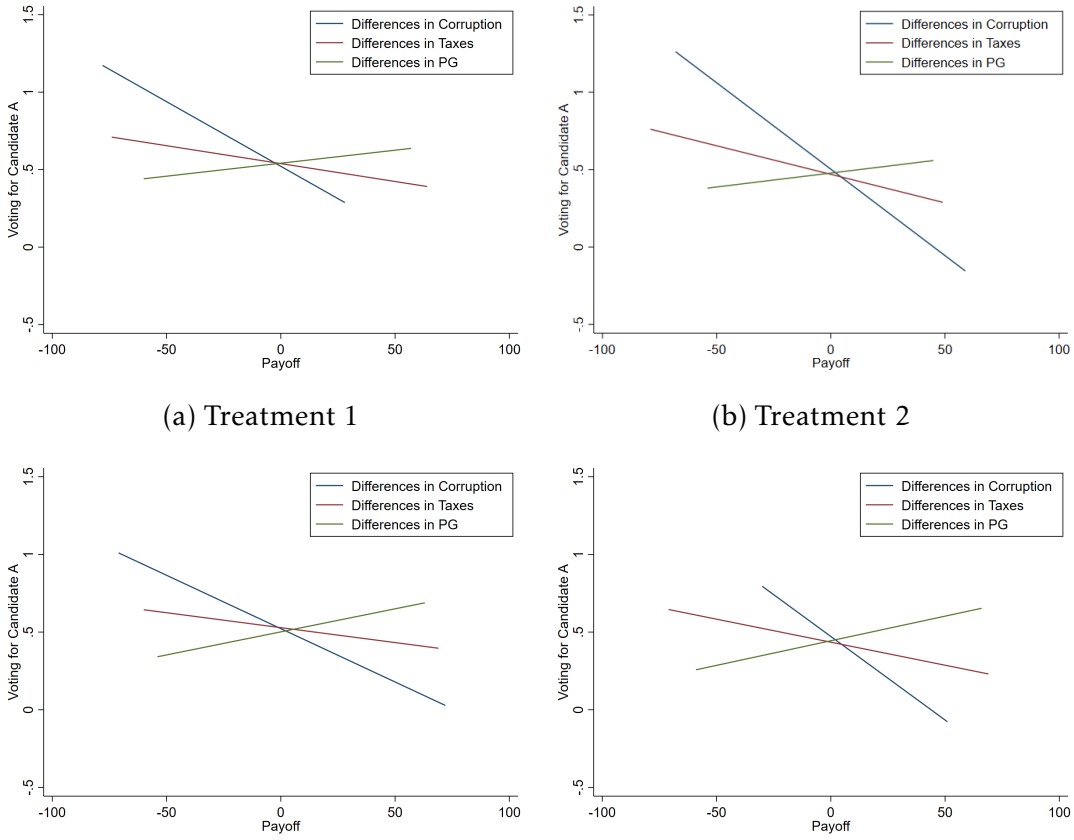
cally significant differences in voter behavior between T1 and T4. The interaction term $T4_g \times \Delta Corruption$ is not significant, suggesting that the introduction of the disproportional power-sharing rule in T4 does not alter voter sensitivity to corruption compared to T1. Similarly, the interaction terms $T4_g \times \Delta Taxes$ and $T4_g \times \Delta Public\ Good$ are also not significant, meaning that voter responses to tax changes and public good provision remain consistent across the two treatments.

Turning to the control variables, partisan bias emerges as a strongly negative and highly significant factor ($p < 0.01$), consistent with its influence in T1–T3. However, its magnitude (−9.977) is slightly smaller than in T1–T3 (−12.24), suggesting that while partisan bias remains a dominant force shaping voter decisions, its impact is somewhat reduced in the context of T1 and T4. Other controls, such as initial income and voter gender, remain non-significant, indicating no systematic influence of these variables on voting behavior.

In conclusion, the findings for T1 and T4 echo the patterns observed in T1–T3, with voters consistently punishing candidates for higher corruption and taxes while

rewarding them for greater public good provision (see Figures 3a–3d for graphical visualization). The introduction of the disproportional power-sharing rule in T4 does not significantly alter voter sensitivities to fiscal policy variables, and the persistent influence of partisan bias continues to shape electoral choices across treatments.

Figure 3: The Impact of Fiscal Policies on Voting for Candidate A.



(a) Treatment 1

(b) Treatment 2

(c) Treatment 3

(d) Treatment 4

A few key differences between the results from T1–T3 and T1–T4 are worth mentioning. First, the effect of corruption (ΔCorruption) on voter decisions is slightly stronger in T1–T4, with coefficients ranging from −0.050 to −0.066, compared to T1–T3, where the range is −0.042 to −0.070. This suggests that voters exhibit heightened sensitivity to corruption under T4's disproportional power-sharing rule.

Another difference is in the interaction terms. In T1–T3, significant interaction effects were observed between automated voters and corruption ($T2_g \times \Delta\text{Corruption}$) and taxes ($T2_g \times \Delta\text{Taxes}$), indicating treatment-specific voter behavior in T2. However, no significant interaction terms are observed in T1–T4, reflecting a lack of distinct treatment-specific differences in voter responses between T1 and T4.

Public good provision (ΔPublic Good) continues to have a significant positive effect

on voter support in both T1–T3 and T1–T4. However, the coefficients in T1–T4 (0.054 to 0.066) are slightly larger than those in T1–T3, suggesting that voters in T4 place somewhat greater emphasis on public good provision.

Finally, partisan bias exhibits subtle differences across the two comparisons. The effect is slightly smaller in T1–T4 (-9.977) than in T1–T3 (-12.24), suggesting reduced polarization or a weaker influence of ideological alignment under T4's disproportional power-sharing rule. This variation invites a closer examination of how partisan bias interacts with voter responsiveness to fiscal policies in different institutional contexts.

7 Conclusion

This paper aimed to explore how corruption, taxes, and the provision of public goods collectively influence electoral accountability. By combining a theoretical model of probabilistic voting with a controlled laboratory experiment, we sought to elucidate the institutional and behavioral conditions under which voters can effectively discipline corrupt or self-serving politicians.

Our key results highlight that *disproportional* power-sharing rules diminish corruption and increase public-good provision, as the higher stakes of electoral victory encourage more voter-friendly policies. Nevertheless, corruption levels remain above the model's benchmark, illustrating the practical constraints imposed by incomplete learning, partisan loyalties, and other behavioural factors. In addition, we find that voters do punish higher corruption but often insufficiently, especially if tax rates or public-good allocations resonate with their interests or partisan predispositions.

From a policy perspective, these findings reinforce that curbing corruption requires more than stricter legal enforcement. It also entails designing electoral institutions that heighten competition and raise the costs of voter alienation. While enhanced transparency can encourage higher public good provision, our results suggest it is not always enough to eliminate rent-seeking behavior. However, limitations such as the finite number of experimental rounds and the simplified lab environment mean that our conclusions should be interpreted with caution.

Future research could explore richer information structures, longer time horizons, or field settings where the salience of corruption may differ. Furthermore, experiments of this nature would benefit from replication in a variety of cultural contexts, since norms of accountability and attitudes toward corruption can be deeply embedded in social and historical factors that our model does not explicitly capture. Pursuing these avenues would deepen our understanding of how democratic processes can

more reliably deliver accountability and social welfare.

A Missing Proofs

A.1 Proof of Proposition 1

To establish that the game $\mathcal{G} = (X, \Pi_C)_{C=A,B}$ possesses a unique pure-strategy equilibrium, we invoke Debreu-Glicksberg-Fan's existence result. We begin observing that the strategy space X is non-empty, compact, and convex. Next, we turn our attention to the payoff functions $\Pi_C(\mathbf{x}_A, \mathbf{x}_B)$. Each of these functions is continuous for all $(\mathbf{x}_A, \mathbf{x}_B) \in X^2$. Finally, we verify that under condition $\mathbb{C}$, each conditional payoff function $\Pi_C(\,\cdot, \mathbf{x}_{-C})$ exhibits strict quasi-concavity on X. This property ensures that each player has a unique best response to the strategies of the other players, which in turn guarantees the uniqueness of the equilibrium.

Consider an arbitrary policy $\overline{\mathbf{x}}_B \in X$, and the resulting conditional payoff $\Pi_A(\,\cdot, \overline{\mathbf{x}}_B)$ for party A. A similar argument can be made for party B. Our goal is to establish that $\Pi_A(\,\cdot, \overline{\mathbf{x}}_B)$ is strictly quasi-concave on X. To this end, it is sufficient to demonstrate that the logarithm of $\Pi_A(\,\cdot, \overline{\mathbf{x}}_B)$ is strictly concave. Note that the second term of $\log(\Pi_A(\,\cdot, \overline{\mathbf{x}}_B))$, specifically, $\log((1-\gamma)\, r_A + R)$, exhibits strict concavity with respect to the party's own strategy r_A. Hence, it rests to show that the first term $\log(\rho(\mathbf{x}^A, \overline{\mathbf{x}}^B))$ is strictly concave in $\mathbf{x}^A$.

To do that, we show that under condition $\mathbb{C}$, the vote and power share functions are both strictly concave. Since the proof is similar in both cases, we only prove it for the power share function. Consider the Hessian associated to $\rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)$, i.e., $H(\mathbf{x}_A) = \left[\frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_j}\right]_{x_i, x_j \in \{t,g,r\}}$. Using the notation of Section 4, for each $\mathbf{x}_A \in X$, $\frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_j} = \rho''(s(\mathbf{x}_A, \overline{\mathbf{x}}_B)) \frac{\partial s(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i} \frac{\partial s(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_j} + \rho'(s(\mathbf{x}_A, \overline{\mathbf{x}}_B)) \frac{\partial^2 s(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_j}$ can be rewritten as $\frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_j} = \frac{\partial s(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i} \frac{\partial s(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_j} \rho'(s(\mathbf{x}_A, \overline{\mathbf{x}}_B)) \Big(\delta(s(\mathbf{x}_A, \overline{\mathbf{x}}_B)) + \delta_{ij}(\mathbf{x}_A, \overline{\mathbf{x}}_B) \Big)$. Therefore, given that by definition $\rho' > 0$, it follows that

$$
\begin{aligned}
\frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_i} < 0 \quad &\Longleftrightarrow \quad \delta(\mathbf{x}_A) < -\Delta^1(\mathbf{x}_A) \\
\begin{vmatrix} \frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_i} & \frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_i \partial x_j} \\ \frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_j \partial x_i} & \frac{\partial^2 \rho(\mathbf{x}_A, \overline{\mathbf{x}}_B)}{\partial x_j \partial x_j} \end{vmatrix} > 0 \quad &\Longleftrightarrow \quad \Delta^2_{[1, \delta(\mathbf{x}_A)]} + \Delta^2_{[2, \delta(\mathbf{x}_A)]} > -\Delta^2(\mathbf{x}_A), \text{ and} \\
|H(\mathbf{x}_A)| < 0 \quad &\Longleftrightarrow \quad \sum_{j \leq 3} \Delta^3_{[j, \delta(\mathbf{x}^A)]}(\mathbf{x}_A) < -\Delta^3(\mathbf{x}_A),
\end{aligned}
$$

which shows by condition $\mathbb{C}$ that the Hessian of $\rho(\,\cdot, \overline{\mathbf{x}}_B)$ is negative definite on X. $\square$

A.2 Proof of Proposition 2

We begin by establishing the equilibrium's symmetry, represented as $\mathbf{x}_A = \mathbf{x}_B$. This symmetry emerges from the identical optimization problems encountered by parties A and B when conditioned on the other party's policy platform. To be precise, let $\mathbf{x}_B \in X$ be fixed, and consider party A's maximization problem. Restricting our analysis to $\mathbf{x}_A = (t_A, g_A, r_A) \in \text{int}(X)$, the first-order conditions for party A are as follows:

$$\begin{cases} \rho\,(1-\gamma) - \lambda_A = 0, \\ ((1-\gamma)r_A + R)\,\rho'(s) \sum\limits_{i\in N} f_i(d_i)\,\frac{\partial V_i(t_A, g_A, r_A)}{\partial g_A} - \lambda_A = 0, \\ ((1-\gamma)r_A + R)\,\rho'(s) \sum\limits_{i\in N} f_i(d_i)\,\frac{\partial V_i(t_A, g_A, r_A)}{\partial t_A} + \lambda_A\, y = 0, \\ \lambda_A\,(t_A y - r_A - g_A) = 0, \end{cases} \tag{10}$$

where λ_A is the Lagrange multiplier associated to the budget constraint, and $d_i = V_i(\mathbf{x}_A) - V_i(\mathbf{x}_B)$. Deriving the first-order conditions for party B and combining it with (10), we have that:

$$\frac{\lambda_A}{\lambda_B} = \frac{\rho}{(1-\rho)}, \tag{11}$$

$$\frac{\lambda_A}{\lambda_B} = \frac{((1-\gamma)r_A + R)\,\rho'(s) \sum\limits_{i\in N} f_i(d_i)\,\partial V_i(\mathbf{x}_A)/\partial g_A}{((1-\gamma)r_B + R)\,\rho'(s) \sum\limits_{i\in N} f_i(d_i)\,\partial V_i(\mathbf{x}_B)/\partial g_B}, \tag{12}$$

$$\frac{\lambda_A}{\lambda_B} = \frac{((1-\gamma)r_A + R)\,\rho'(s) \sum\limits_{i\in N} f_i(d_i)\,\partial V_i(\mathbf{x}_A)/\partial t_A}{((1-\gamma)r_B + R)\rho'(s) \sum\limits_{i\in N} f_i(d_i)\,\partial V_i(\mathbf{x}_B)/\partial t_B}. \tag{13}$$

Suppose, by contradiction, that the equilibrium is not symmetric, that is, $\mathbf{x}_A \neq \mathbf{x}_B$. Without loss of generality, assume that $g_A > g_B$. From (12) and (13), we can infer that $t_A < t_B$, since $\partial V_i(\mathbf{x})/\partial g\,(> 0)$ is decreasing in g, and $\partial V_i(\mathbf{x})/\partial t\,(< 0)$ is decreasing in t. Moreover, since the budget constraint is binding, $r_A < r_B$; and using (12), $g_A > g_B$ and $r_A < r_B$ imply that $\frac{\lambda_A}{\lambda_B} < 1$. In addition, notice that $s(\mathbf{x}_A, \mathbf{x}_B) > 1/2$, and consequently $\rho(\mathbf{x}_A, \mathbf{x}_B) > 1/2$. Hence, from (11), it follows that $\frac{\lambda_A}{\lambda_B} > 1$, a contradiction. Therefore, equilibrium policy $\mathbf{x}_C = \mathbf{x}^*$ for each $C = A, B$.

To characterize such policy, notice that $s(\mathbf{x}^*, \mathbf{x}^*) = \rho(\mathbf{x}^*, \mathbf{x}^*) = 1/2$, with $\frac{\partial \rho}{\partial s} = \eta$; and from (11) and the first equation of (10), $\lambda_A = \lambda_B = \lambda^* = (1-\gamma)/2 > 0$. Then, using the budget constraint, $t^* = \frac{r^* + g^*}{y}$. Finally, the optimal values of g and r are derived from

the remaining two equations of (10), namely,

$$((1-\gamma)r^* + R)\,\eta\,\frac{f(0)}{n}\sum_{i\in N}\left[a_i'(g^*) - u_i'\left(\left(1-\frac{r^*+g^*}{y}\right)y_i\right)\frac{y_i}{y}\right] = 0,$$

and

$$\frac{1-\gamma}{2} - ((1-\gamma)r^* + R)\,\eta\,\frac{f(0)}{n}\sum_{i\in N}\left[u_i'\left(\left(1-\frac{r^*+g^*}{y}\right)y_i\right)\frac{y_i}{y} + \beta_i\right] = 0 \qquad \square$$

A.3 Proof of Proposition 3

The comparative statics results displayed in Proposition 3 are derived from the system of first-order conditions (10), using implicit function theorem. To be precise, let

$$\begin{aligned}
F_1(r,g,\eta,R,\gamma,\beta_i,f(0)) &= ((1-\gamma)r + R)\,\eta\,\frac{f(0)}{n}\sum_{i\in N}\left[a_i'(g) - u_i'\left(\left(1-\frac{r+g}{y}\right)y_i\right)\frac{y_i}{y}\right],\\
F_2(r,g,\eta,R,\gamma,\beta_i,f(0)) &= \frac{1-\gamma}{2} - ((1-\gamma)r + R)\,\eta\,\frac{f(0)}{n}\sum_{i\in N}\left[u_i'\left(\left(1-\frac{r+g}{y}\right)y_i\right)\frac{y_i}{y} + \beta_i\right].
\end{aligned}$$

For each $k = 1, 2$, and each $x \in \{r, g, \eta, R, \gamma, \beta_i, f(0)\}$, denote by $F_{kx}\left(=\frac{\partial F_k}{\partial x}\right)$ the first-order partial derivatives of F_k with respect to x. From the implicit function theorem, we know that:

$$\begin{pmatrix}\frac{\partial g}{\partial \eta}\\ \frac{\partial r}{\partial \eta}\end{pmatrix} = -\begin{bmatrix}F_{1g} & F_{1r}\\ F_{2g} & F_{2r}\end{bmatrix}^{-1}\cdot\begin{pmatrix}F_{1\eta}\\ F_{2\eta}\end{pmatrix}.$$

Simple algebraic manipulations reveal that:

$$\frac{\partial g}{\partial \eta} = -\frac{F_{2r}\cdot F_{1\eta} - F_{1r}\cdot F_{2\eta}}{\det(M)}, \tag{14}$$

$$\frac{\partial r}{\partial \eta} = -\frac{-F_{2g}\cdot F_{1\eta} + F_{1g}\cdot F_{2\eta}}{\det(M)}, \tag{15}$$

$$\frac{\partial t}{\partial \eta} = -\frac{F_{1\eta}\cdot(F_{2r} - F_{2g}) + F_{2\eta}\cdot(F_{1g} - F_{1r})}{\det(M)}, \tag{16}$$

where the matrix $M = \begin{bmatrix}F_{1g} & F_{1r}\\ F_{2g} & F_{2r}\end{bmatrix}$. Notice that since each voter's preferences on X are strictly concave, $\det(M) > 0$. Moreover, it is also easy to prove that $F_{1g} < 0$, $F_{1\eta} = 0$, $F_{1r} < 0$, $F_{2\eta} < 0$, $F_{2r} < 0$, $F_{2g} < 0$, and $F_{1g} < F_{1r}$. Thus, from (14)–(16), it follows that $\frac{\partial r}{\partial \eta} < 0$, $\frac{\partial g}{\partial \eta} > 0$, and $\frac{\partial t}{\partial \eta} < 0$.

The rest of the comparative statics relationships are established using a similar

reasoning, and the fact that $F_{1x} = 0$ for each $x \in \{\gamma, R, \beta_i, f(0)\}$, and that $F_{2x} < 0$ for each $x \in \{\gamma, R, \beta_i, f(0)\}$. □

A.4 Proof of Proposition 4

The utilitarian social planner aims to maximise the social welfare function $W(t, g, r) = \sum_{i \in N} V_i(t, g, r) = \sum_{i \in N} [a_i(g) + u_i((1-t)y_i) - \beta_i r]$ subject to the budget constraint $g + r \leq ty$. Note that W is a continuous function and the feasible choice set is non-empty and compact. Therefore, by Weierstrass theorem, W attains a maximum $(\hat{t}, \hat{g}, \hat{r})$ on X. To characterise this maximum, we substitute the budget constraint into W and we differentiate with respect to g and r, yielding the following two equations:

$$\frac{\partial W(t, g, r)}{\partial g} = \sum_{i \in N} a_i'(g) - \sum_{i \in N} u_i'((1-t)y_i)\frac{y_i}{y}, \tag{17}$$

$$\frac{\partial W(t, g, r)}{\partial r} = -\left(\sum_{i \in N} u_i'((1-t)y_i)\frac{y_i}{y} + \sum_{i \in N} \beta_i\right). \tag{18}$$

Therefore, at the planner's social optimum $(\hat{t}, \hat{g}, \hat{r})$, equations (17) and (18) imply that: (i) the optimal value of political rents $\hat{r}$ is equal to zero, and (ii) the public good $\hat{g}$ satisfies the condition $\sum_{i \in N} a_i'(\hat{g}) - \sum_{i \in N} u_i'\left(\left(1 - \hat{t}\right) y_i\right) \frac{y_i}{y} = 0$, where the tax rate $\hat{t} = \frac{\hat{g}}{y}$ is determined by the budget constraint. This concludes our proof. □

A.5 Proof of Proposition 5

The proof proceeds as follows. By Proposition 4, the socially optimal policy $\hat{\mathbf{x}} = (\hat{t}, \hat{g}, \hat{r})$ is independent of the model parameters η, R, γ, β_i, and $f(0)$. Consequently, the welfare loss $L(\hat{\mathbf{x}}, \mathbf{x}^*)$ is influenced by these parameters through the political equilibrium $\mathbf{x}^* = (t^*, g^*, r^*)$. To analyse these effects, we totally differentiate $L(\hat{\mathbf{x}}, \mathbf{x}^*)$ with respect to each $\mu \in \{\eta, R, \gamma, \beta_i, f(0)\}$, yielding the following equation:

$$\frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial \mu} = \frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial g^*} \cdot \frac{\partial g^*}{\partial \mu} + \frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial r^*} \cdot \frac{\partial r^*}{\partial \mu}.$$

Note that

$$\frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial g^*} = \sum_{i \in N} \left(u_i' \left(\left(1 - \frac{r^* + g^*}{y} \right) y_i \right) \frac{y_i}{y} - a_i'(g^*) \right) = 0$$

by equation (2), and

$$\frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial r^*} = \sum_{i \in N} \left(u_i' \left(\left(1 - \frac{r^* + g^*}{y} \right) y_i \right) \frac{y_i}{y} + \beta_i \right) > 0$$

by equation (3). Furthermore, from Proposition 3, we have $\frac{\partial r^*}{\partial \mu} < 0$. Therefore,

$$\frac{\partial L(\hat{\mathbf{x}}, \mathbf{x}^*)}{\partial \mu} < 0 \quad \text{for all} \quad \mu \in \{\eta, R, \gamma, \beta_i, f(0)\}. \quad \square$$

References

Abbink, K. (2006): "14 Laboratory experiments on corruption," *International handbook on the economics of corruption*, 418.

Afridi, F., S. Bhattacharya, A. Dhillon, and E. Solan (2024): "Electoral Competition, Electoral Uncertainty and Corruption: Theory and Evidence from India," *Journal of Economic Behavior and Organization*.

Anduiza, E., A. Gallego, and J. Muñoz (2013): "Turning a blind eye: Experimental evidence of partisan bias in attitudes toward corruption," *Comparative Political Studies*, 46, 1664–1692.

Arias, E., P. Balán, H. Larreguy, J. Marshall, and P. Querubín (2019): "Information provision, voter coordination, and electoral accountability: Evidence from Mexican social networks," *American Political Science Review*, 113, 475–498.

Ashworth, S., E. B. De Mesquita, and A. Friedenberg (2017): "Accountability and information in elections," *American Economic Journal: Microeconomics*, 9, 95–138.

Besley, T. and A. Case (1995): "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits," *The Quarterly Journal of Economics*, 110, 769–798.

Besley, T., T. Persson, and D. M. Sturm (2010): "Political competition, policy and growth: theory and evidence from the US," *The Review of Economic Studies*, 77, 1329–1352.

Bobkova, N. and H. Egbert (2012): "Corruption investigated in the lab: A survey of the experimental literature," *International Journal of Latest Trends in Finance and Economic Sciences*, 337–349.

Bol, D., K. Matakos, O. Troumpounis, and D. Xefteris (2019): "Electoral rules, strategic entry and polarization," *Journal of Public Economics*, 178, 104065.

Bruns, C. and O. Himmler (2016): "Mass media, instrumental information, and electoral accountability," *Journal of Public Economics*, 134, 75–84.

Chang, E. C. and M. A. Golden (2007): "Electoral systems, district magnitude and corruption," *British journal of political science*, 37, 115–137.

Cubel, M., A. Papadopoulou, and S. Sánchez-Pagés (2024): "Identity and political corruption: a laboratory experiment," *Economic Theory*, 1–24.

Devdariani, S. and A. V. Hirsch (2023): "Voter attention and electoral accountability," *Journal of Public Economics*, 224, 104918.

Duggan, J. and C. Martinelli (2020): "Electoral accountability and responsive democracy," *The Economic Journal*, 130, 675–715.

Eggers, A. C., N. Vivyan, and M. Wagner (2018): "Corruption, accountability, and gender: Do female politicians face higher standards in public life?" *The Journal of Politics*, 80, 321–326.

Eggers, A. C. et al. (2014): "Partisanship and electoral accountability: Evidence from the UK expenses scandal," *Quarterly Journal of Political Science*, 9, 441–472.

Ferraz, C. and F. Finan (2011): "Electoral accountability and corruption: Evidence from the audits of local governments," *American Economic Review*, 101, 1274–1311.

Foucault, M., K. Seki, and G. D. Whitten (2017): "Good times, bad times: Taxation and electoral accountability," *Electoral Studies*, 45, 191–200.

Garfinkel, M. and S. Skaperdas (2007): "Economics of conflict: An overview," in *Handbook of Defense Economics*, ed. by T. Sandler and K. Hartley, North-Holland, vol. 2, 649–709.

Golden, M. A. and P. Mahdavi (2015): "The institutional components of political corruption 1," *Routledge handbook of comparative political institutions*, 404–420.

Heggedal, T.-R., L. Helland, and R. Morton (2022): "Can paying politicians well reduce corruption? The effects of wages and uncertainty on electoral competition," *Games and Economic Behavior*, 135, 60–73.

Hellwig, T. and D. Samuels (2008): "Electoral accountability and the variety of democratic regimes," *British Journal of Political Science*, 38, 65–90.

Herrera, H., M. Morelli, and S. Nunnari (2016): "Turnout across democracies," *American Journal of Political Science*, 60, 607–624.

Hirshleifer, J. (1989): "Conflict and rent-seeking success functions: Ratio vs. difference models of relative success," *Public choice*, 63, 101–112.

Isaac, R. M. and D. A. Norton (2012): "New advances in experimental research on corruption," .

Jones, D. R. (2010): "Partisan polarization and congressional accountability in house elections," *American Journal of Political Science*, 54, 323–337.

Kunicová, J. (2006): "Democratic institutions and corruption: incentives and constraints in politics," *International handbook on the economics of corruption*, 1, 140.

Kunicova, J. and S. Rose-Ackerman (2005): "Electoral rules and constitutional structures as constraints on corruption," *British journal of political science*, 35, 573–606.

Lindbeck, A. and J. W. Weibull (1987): "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, 52, 273–297.

Lowry, R. C., J. E. Alt, and K. E. Ferree (1998): "Fiscal policy outcomes and electoral accountability in American states," *American political science review*, 92, 759–774.

Martin, L. and P. J. Raffler (2021): "Fault lines: the effects of bureaucratic power on electoral accountability," *American Journal of Political Science*, 65, 210–224.

Martinelli, C. (2022): "Accountability and grand corruption," *American Economic Journal: Microeconomics*, 14, 645–679.

Martinelli, C. and N. C. S. Esandi (2022): "Career incentives and corruption: A lab experiment," *Available at SSRN 4231320*.

Matakos, K., O. Troumpounis, and D. Xefteris (2016): "Electoral rule disproportionality and platform polarization," *American Journal of Political Science*, 60, 1026–1043.

Min, B. (2008): "Democracy and light: Electoral accountability and the provision of public goods," in *annual meeting of the Midwest Political Science Association*.

Olken, B. A. (2009): "Corruption perceptions vs. corruption reality," *Journal of Public economics*, 93, 950–964.

Persson, T. and G. Tabellini (2000): *Political economics: explaining economic policy*, MIT press.

Persson, T., G. Tabellini, and F. Trebbi (2003): "Electoral rules and corruption," *journal of the European Economic Association*, 1, 958–989.

Polo, M. (1998): "Electoral competition and political rents," .

Rose, J. (2014): "Corruption and the problem of perception," in *Routledge handbook of political corruption*, Routledge, 172–182.

Rudolph, L. and T. Däubler (2016): "Holding individual representatives accountable: The role of electoral systems," *The Journal of Politics*, 78, 746–762.

Saporiti, A. (2014): "Power sharing and electoral equilibrium," *Economic Theory*, 55, 705–729.

Solaz, H., C. E. De Vries, and R. A. De Geus (2019): "In-group loyalty and the punishment of corruption," *Comparative Political Studies*, 52, 896–926.

Treisman, D. (2014): "What does cross-national empirical research reveal about the causes of corruption?" *Routledge handbook of political corruption*, 95–109.

Tullock, G. (1980): "Efficient rent seeking," in *Toward a Theory of the Rent-Seeking Society*, ed. by J. Buchanan, R. Tollison, and G. Tullock, College Station: Texas AM University Press, 97–112.

Zudenkova, G. (2018): "Electoral accountability and interdistrict competition," *American Economic Journal: Microeconomics*, 10, 143–176.