

Non-Electoral Accountability: Citizen Sanctions on Traditional Leaders in Sierra Leone

November 2, 2025

Word count: 4993

Abstract

How do citizens in Sierra Leone perceive the legitimacy and feasibility of sanctioning their chiefs outside of elections? This study investigates perceptions of non-electoral sanctions through a pre-registered survey experiment in Sierra Leone. We find that citizens view indirect sanctions—such as appealing to higher authorities—as more legitimate and feasible than direct actions, and that the range of acceptable sanctions expands with the severity of the offense. Community elders' involvement increases the perceived legitimacy of sanctions, highlighting their role as political intermediaries. Finally, respondents' social status moderates their perceptions of both the legitimacy and the feasibility of sanctions. These results suggest that even in highly hierarchical settings, citizens may retain some capacity to discipline chiefs, though accountability seems primarily mediated through vertical institutions rather than direct collective action.

Roughly 30% of the global population lives under customary laws and traditional governance structures, highlighting their widespread presence across continents (Baldwin & Holzinger 2019). Although traditional leaders—particularly village chiefs—are not subject to regular elections, they often enhance government responsiveness, facilitate collective action, and broker resources (Baldwin 2016; Honig 2017; Murtazashvili & Murtazashvili 2016). One explanation is that village members possess the capacity to sanction their leaders, which may compel those leaders to act in the community’s interest. In particular, non-electoral sanctions have emerged as a prominent theoretical mechanism through which traditional leaders are held accountable (Carlson 2021; Mattingly 2016; Paller 2014; Wilfahrt 2018).

This study builds on and contributes to existing research by formally testing the perceived legitimacy and feasibility of non-electoral sanctions through a novel survey experiment¹. It examines whether, and under what conditions, village members view sanctioning their chiefs as a justified response when leaders act against the interests of the community². Scholars highlight two main mechanisms through which citizens may influence local leaders outside of electoral institutions. First, research emphasizes that leaders who are socially and economically embedded in their communities may be more responsive to citizen interests (Baldwin 2016). In such contexts, villagers may withhold cooperation, voice public criticism, or even threaten non-compliant chiefs (Baldwin 2016; Chen et al. 2016; Tsai 2007; Zhong & Zeng 2024). Second, scholars have pointed to the role of hierarchical political institutions. Honig (2022), for ex-

¹ This project was approved by the Office of Sierra Leone Ethics and Scientific Review Committee (SLESRC n°020/04/2023). A Pre-Analysis Plan for this project has been registered with OSF: <https://osf.io/8r7zm>.

² Sierra Leone is consistently ranked as one of the most corrupt countries in the world, with frequent allegations of nepotism. The misappropriation of community funds by low-level chiefs is common. One such instance is provided by Anti-Corruption Commission of Sierra Leone (2024). To facilitate understanding of the experimental scenario, and to use a common case of chiefs’ behavior that does not represent community interests, we focus on the misappropriation of community funds for the chief’s private benefit. The paper’s findings, therefore, do not necessarily translate to cases where chiefs commit different offenses, as the sanctioning norms and what may be perceived as legitimate and feasible sanctions may differ depending on the type of norm transgression.

ample, shows that hierarchy can generate horizontal accountability among chiefs by creating incentives for compliance with higher-level authorities. Yet, as Baldwin (2025) notes, little empirical work has examined whether such structures also enable downward accountability to citizens. Building on this gap, we theorize that villagers have two channels through which they may sanction chiefs: either by directly enforcing social norms themselves, or by appealing to higher authorities who can sanction local leaders. We make a first step toward testing this theoretical claim by examining whether citizens view these sanctions as legitimate and feasible. The study was conducted in Sierra Leone in April and May 2023, following pre-registered hypotheses (see Appendix C for the pre-analysis plan). In this context, chiefs wield considerable authority, face limited political competition, and have been described as capturing civil society (Acemoglu et al. 2014). Scholars often characterize the institutional environment as lacking robust constraints on chiefs, with weak formal mechanisms of accountability. At the same time, village chiefs are deeply embedded in the local social and political fabric, making Sierra Leone a particularly relevant setting for examining informal mechanisms of accountability.

We conducted a survey experiment to test whether villagers would endorse sanctioning their village chief if the chief were hypothetically to steal from a community project. The experiment also examined the role of political intermediaries—specifically, community elders—as mediators shaping sanctioning preferences. Our findings show that villagers primarily support indirect sanctions, such as reporting to higher authorities. In contrast, many direct sanctions are not widely perceived as legitimate forms of political behavior. Those preferences differ substantially by gender, income, and voting rights, with men, lower-income respondents, and those with voting rights viewing, on average, a wider set of sanctions as legitimate and feasible. Notably, when community elders publicly criticize the chief, the range of acceptable sanctions expands, underscoring their role in coordinating local accountability and reinforcing social norms.

This study makes three contributions. First, it broadens our theoretical understanding of accountability mechanisms beyond electoral settings, complementing classic selectorate models of accountability (Balasuriya 2023; Mesquita et al. 2005) and advancing discussions on the sources of chiefs' authority (Baldwin & Ricart-Huguet 2022). Categorizing sanctioning chan-

nels as direct or indirect offers a valuable analytical framework. Second, our study clarifies how chiefs in Sierra Leone may be held accountable by their communities (Acemoglu et al. 2014; Baldwin & Holzinger 2019), identifying the political tools available to citizens to hold leaders accountable. We show that indirect sanctions are generally seen as more legitimate and less costly than direct sanctions. Third, this research contributes to the growing literature on political intermediaries, with particular attention to community elders (Baldwin et al. 2022). We demonstrate that community elders influence the range of sanctions considered as legitimate and feasible by village members.

Institutional context

Sierra Leone's chieftaincy system features strong vertical authority and limited formal accountability, patterns common in rural sub-Saharan Africa (Baldwin 2025). The study took place in Southern Province's Bonthe and Moyamba districts, where chiefs hold substantial power and checks are limited. It covers 12 chiefdoms and 77 villages. Chiefdoms, led by paramount chiefs elected by tribal authorities, are key governance units (Appendix B for further information). Each chiefdom is divided into sections governed by section chiefs above village chiefs. Chiefs act as state clients and community patrons, mediating between rural populations and national authorities (Becorpi 2018). They allocate resources, sustain patronage networks, and mobilize electoral support in exchange for state-backed resources (Becorpi 2018; Labonte 2012). Though formally part of the state, chiefs sometimes compete with government institutions (Grieco 2024; Henn 2022). Despite 2004 decentralization reforms, many rural residents prefer traditional leaders, who are seen as better defenders of customary land rights and more effective dispute resolvers (Fanthorpe 2006; Sawyer 2008). Village chiefs remain trusted conflict authorities despite concerns over corruption and nepotism (Ruppel & Leib 2022; Sawyer 2008).

Accountability at the village level is complex. Town chiefs, though formally elected by taxpaying members of landowning lineages, serve for decades and often for life (Bulte et al. 2018). In our sample, the last elections for town chiefs ranged from 1982 to 2023, averaging around

2012—over ten years before data collection in April–May 2023. Such infrequency makes non-electoral sanctions particularly important. Yet, in Sierra Leone, the chieftaincy has been described as a colonial invention designed to make chiefs accountable upward rather than to their communities (Acemoglu et al. 2014; Labonte 2012; Mihaylova 2023). Consistent with this, about 24% of villages in our sample have chiefs who were elected after their predecessors were suspended, reflecting the frequency of higher-level interventions, which Bulte et al. (2018) argue are often strategic and self-serving.³ Still, Voors et al. (2018) argue that communities can sometimes leverage these vertical ties to discipline their town chiefs.

Downward accountability is further limited by the exclusionary structure of rural politics. Only members of “chiefly families”—sometimes as few as 13% of households—are eligible for office (Labonte 2012). Labonte (2012) also notes that, while challenging chiefs through public forums is theoretically possible, “non-elites are risk-averse in claiming their rights from elites, airing grievances in public forums, or demanding accountability.” This exclusion has historically fueled grievances and contributed to civil war (Humphreys & Weinstein 2006; Mokuwa et al. 2011; Richards 2005), echoing Conteh (2013) mentions of the use of revolts and murders in extreme cases. Yet, paradoxically, satisfaction with chiefs is high: Casey et al. (2012) report that 94% of households expressed approval in a comparable sample of villages. To date, however, there has been no systematic assessment of whether villagers can effectively discipline chiefs or mobilize higher-level authorities. This paper aims to fill this gap by empirically testing whether community members view a range of sanctions as legitimate and feasible, particularly when chiefs misappropriate community project resources, a common form of theft in Sierra Leone.

Finally, in 87% of villages, chiefs are supported by councils of elders, heads of major descent groups (Leach 2022), who claim legitimacy as early settlers and control land rights. Elders advise on land allocation and coordinate with chiefs. Village chiefs in our sample are well embedded: 77% of respondents were born in the village, 73% of chiefs own farms there, and

³ We did not collect information on the reasons for these suspensions.

only 21% have formal jobs⁴. Appendix B provides additional descriptive statistics on the institutional context.

Theory and hypotheses

Accountability requires that leaders adhere to established behavioral standards, with the risk of sanctions for failing to do so (Chen et al. 2016; Grant & Keohane 2005). At its core, accountability serves to penalize the illegitimate use of power and rests on two foundational components: access to information and the capacity to impose sanctions (Grant & Keohane 2005). While access to information is a necessary precondition, this paper focuses on the second condition—the capacity to impose sanctions.

In electoral democracies, accountability is often exercised through periodic, competitive elections that allow citizens to select and remove leaders. In the absence of regular elections, however, electoral accountability is absent or weak (Baldwin 2025). Instead, community members must rely on non-electoral accountability mechanisms (Baldwin 2016, 2025). We focus here on the repertoire of non-electoral sanctions available to community members to sanction their village leaders when it does not behave in the community interests. Our theoretical framework is further illustrated in Appendix A.

Community interests refer to collective welfare and adherence to established social norms, including equitable allocation of resources, fair use of customary taxes, provision of public goods, protection of vulnerable community members, and adherence to local moral expectations. We focus specifically on one illustrative instance: the management of a community development project in which funds are misappropriated or stolen, a common issue in Sierra Leone (Anti-Corruption Commission of Sierra Leone 2024). Chiefs are expected to advance community interests in such projects—for example, by ensuring that funds are used for the intended public goods, reporting transparently to community members, and allocating resources equitably.

Sanctions are actions that impose costs—or withhold benefits—from a leader following a perceived violation (Meng et al. 2023; Ostrom 2005). Sanctions are not reducible to actions alone:

⁴ Data on residency and time spent in villages were not collected.

their effectiveness depends on the social conditions under which they can be mobilized (Ostrom 2000). We distinguish three interrelated but conceptually distinct dimensions of sanctioning capacity. First, sanction norms refer to the socially prescribed scripts that dictate what kinds of responses are considered appropriate following a violation (Ostrom 2005). They are shared expectations about how community members should respond when leaders fail in their duties (Ostrom 2005). Second, legitimacy concerns whether a given sanction is broadly recognized as rightful and appropriate within the community. A sanction norm may exist, but if citizens perceive it as biased, disproportionate, or captured by particular groups, its legitimacy is undermined and enforcement becomes unlikely. Third, feasibility refers to the practical ability to carry out a sanction. Even when a sanction is normatively prescribed and considered legitimate, it may not be feasible if it requires resources, coordination, or authority beyond the citizens' reach. Sanctioning capacity thus involves: (1) the existence of norms that render sanctions socially meaningful, (2) legitimacy that makes them acceptable, and (3) feasibility that makes them practically implementable and enforceable. In this study we focus on legitimacy and feasibility as two necessary but not sufficient conditions for non-electoral accountability.

Finally, we categorize the repertoire of sanctions according to their pathway of enforcement. Direct sanctions are those citizens can impose themselves. Chiefs who are embedded in their communities are more likely to provide public goods, as their interests align with those of their constituents (Baldwin 2016), and they are motivated to earn the moral standing of citizens (Baldwin 2025; Tsai 2007). Anthropological studies document various citizen strategies to pressure leaders (Arnall et al. 2013), echoing Scott's "weapons of the weak" (Scott 1985), but the full scope of legitimate sanctions remains underexplored. Potential direct sanctions include public blames (Labonte 2012) or restricting chiefs' access to communal institutions. In farming and fishing communities, collective labor depends on trust (Bulte et al. 2018), so chiefs neglecting communal interests risk losing labor support. Appendix A further discusses the set of sanctions considered in this study and their mentions in the literature.

On the other hand, indirect sanctions operate through intermediaries—community elders or higher-level chiefs such as section and paramount chiefs—who impose sanctions on town chiefs. These mechanisms rely on the vertical structure of traditional authority in Sierra Leone,

where village chiefs are formally subordinate to section and paramount chiefs. This hierarchy is both codified in customary law and broadly recognized within communities, providing institutional oversight: severe violations of community interests by a village chief can lead to suspension or removal. These formal hierarchical ties create an institutional pathway through which citizens may strategically pursue indirect sanctions.

Overall, town chiefs are likely to be more responsive to some sanctioning agents than to others, making the pathway of enforcement critical for understanding accountability. This distinction matters because chiefs are not equally responsive to all forms of pressure: they may be more attentive to sanctions imposed by authoritative intermediaries than to citizen mobilization, as often stated in Sierra Leone (Bulte et al. 2018; Labonte 2012). While sanctions also differ in the distribution of their social costs (some primarily burdening the chief, others imposing collective costs), our focus here is on the enforcement pathway, which conditions accountability dynamics. This framework echoes existing research on accountability in other non-electoral contexts (Anderson et al. 2019; Chen et al. 2016) and complement selectorate models of accountability (Balasuriya 2023). Finally, for downward accountability to be effective, some sanctions require social coordination and, consequently, the resolution of inherent collective action problems. This is another necessary condition, which is beyond the scope of this paper.

Expectations

We theorize that non-electoral sanctions may take two forms: citizens can appeal to higher authorities⁵ to sanction a village chief, or they can directly sanction the chief themselves⁶.

Hypothesis 1: Village members view sanctions against chiefs who violate community interests as legitimate and feasible.

Our pre-registered main hypothesis stated that when undemocratic village leaders do not act in the interest of their communities, citizens sanction them through a variety of social, economic, and political channels, with a preference for lower-cost options (Hypothesis 1; see Appendix G). Although the hypothesis originally concerned sanctioning behavior, our design does not ob-

⁵ such as elders, section and paramount chiefs

⁶ actions such as blaming, threatening, refusing labor, withholding taxes, or withholding marriage ties

serve sanctions directly. Instead, we test whether citizens view different sanctions as legitimate and feasible. We focus primarily on the extensive margin, namely whether citizens endorse any sanction, while also considering the intensive margin of sanction severity as a secondary outcome. By contrast, preferences for lower-cost sanctions are not tested in the main analysis and are only supported with qualitative evidence from five focus groups in Appendix K, so this aspect receives less emphasis.

We also investigate the role of village elites, the council of elders, in driving sanctioning behavior of other community members. The council's sanctions against the chief might serve as a sufficient deterrent, reducing the need for further community action. Prior research has revealed that the pressure exerted by community elders effectively shapes and influences the behavior of village chiefs (Baldwin et al. 2022). Consequently, when community elders have already taken action, the efficacy of community members' pressure is diminished, leading us to expect their abstention from participating in such actions. The wording of pre-registered hypothesis 2 was simplified for clarity; content and predictions remain unchanged (see Appendix G).

Hypothesis 2: community members will be less willing to view sanction as legitimate and feasible if the council of elders already blamed the chief for their action.

Additional hypotheses from the pre-analysis plan are outside this paper's scope and are addressed separately in another project.

Research design

Examining whether and how community members sanction leaders who act against their interests raises methodological challenges. First, the lack of official records on sanctions complicates tracing and analysing such incidents. Second, effective sanctioning mechanisms may deter violations of community interests, making theft and sanctions rare and hard to study. Third, the region's history of civil conflict, tied to the exclusion of local voices (Peters 2011), may discourage political actors from discussing sensitive behaviours, complicating data collection. In such a setting, a survey experiment provides a strong justification because it allows

us to study sanctioning responses to off-equilibrium violations that might never be directly observable. Consequently, we use a novel survey experiment designed to measure individual preferences for non-electoral sanctions.

Communities were randomly selected from villages located near mangrove resources, with eligibility based on population size. This sampling frame was defined by the goals of the broader research project, which focused on forest-dependent livelihoods and community-forest relationships. Although this paper does not analyse forest use directly, selecting such villages ensured that local chiefs were embedded in the community, a key theoretical scope condition for this study. Villages with more than 200 households were excluded, as chiefs in larger communities are often less integrated. Villages with fewer than 20 households were also excluded due to sample size requirements in other parts of the project.

In each village, we randomly selected 12 household heads using a two-step process: (1) a full household listing, and (2) a random sampling.

Household survey experiment⁷

We implemented a vignette survey experiment using five hypothetical scenarios to identify sanctions deemed legitimate by community members and assess how the behavior of community elders shapes sanctioning preferences.

The main scenario reflects a common issue: misallocation of NGO-led development resources, such as cookstoves, agricultural inputs, or mini-grids. Elite capture of such resources is well-documented across sectors like education (Mbiti 2016; Reinikka & Svensson 2002), health (Azfar & Gurgur 2008; Njong & Ngantcha 2013), and development (Carlson & Seim 2020). While difficult to trace, these leakages can be curbed by accountability mechanisms (Anti-Corruption Commission of Sierra Leone 2024; Carlson & Seim 2020). Our vignette features a

⁷ The pre-analysis plan included two scenarios: a community project and a land deal. During fieldwork, the land deal scenario was found to misalign with contextual realities—lands can only be leased, not sold, and transactions often involve the paramount chief. Additionally, the scenario lacked a pure control, limiting causal inference. In line with the pre-analysis plan, results for the land deal are presented in Appendix G.6, and are consistent with the main findings.

chief who appropriates part of an NGO project’s funds. Though many NGOs distribute goods rather than cash,⁸ we simplified the scenario by focusing on funds for clarity. Neutral language was used, and enumerators were instructed not to use the word “steal” in local translations to avoid bias.

Each participant received one randomly assigned treatment. Randomization was conducted independently within each village. We also block-randomized by gender to improve the precision of estimates, given the modest overall sample size (907 respondents). Table 1 outlines these conditions. The control presents a well-managed project. Treatment 1 introduces minor misappropriation⁹. Treatment 4, a larger sum. These gauge citizens’ sanctioning repertoire and whether responses vary by severity. The inclusion of a control arm serves to capture baseline sanctioning attitudes, rather than sanctions triggered by a specific act of misappropriation. Even without evidence of wrongdoing, some respondents considered sanctions legitimate, likely reflecting a general skepticism toward chiefs. This baseline is essential for identifying the marginal effect of misappropriation on sanction preferences. In addition, our design leverages the comparison between Treatment 1 (minor misappropriation) and Treatment 4 (major misappropriation) to examine the intensive margin of sanctioning, ensuring that inferences are not solely dependent on control–treatment contrasts. Treatments 2 and 3 introduce council of elders’ behavior, allowing analysis of how elite dynamics shape preferences.

Appendix E shows balance across 13 covariates. F-tests are insignificant at the 5% level, indicating balance. However, Treatment 1 recipients are slightly less trustful of the chief and less employed than controls; Treatment 2 recipients are older than those in Treatment 3. We control for these imbalances in robustness tests (Appendix J).

⁸ In Sierra Leone, mini-grids, agricultural inputs, and cash transfers often involve chiefs in resource allocation.

⁹ In many rural development contexts, it is common for chiefs to receive compensation for their time when overseeing or implementing community projects run by an NGO. If community members view this as the norm, Treatment 1 may not effectively prime perceptions of misappropriation. However, during focus groups and piloting, it became clear that even a small appropriation by the town chief was not well received by the community.

Appendix F reports comprehension rates: 95% understood the amount taken, and 92% understood the elders' behavior.¹⁰

Table 1: Description of the control and treatment conditions.

Conditions	Details
Control	<i>In a village in Sierra Leone, an NGO developed a project for the development of the community. The town chief played a key role in managing the project at the village level.</i>
Treatment 1	Control + During the project, the town chief took a very small part of the money for his benefit.
Treatment 2	Treatment 1 + The elders in the village went to the chief and sermoned him.
Treatment 3	Treatment 1 + The elders in the village did not sermon the chief.¹¹
Treatment 4	Control + During the project, the town chief took half of the money for his benefit.

Notes:

Enumerators read the scenarios to participants in local languages (Sherbro, Mende, or Krio) using tablets.

After reading the scenario, enumerators asked respondents five questions:

- Outcome 1: Agreement with the chief's behaviour (1–5 scale);
- Outcome 2: Whether citizens should respond (binary);
- Outcome 3: If yes, specify actions (open-ended);¹²
- Outcome 4 (main): Number of sanctions deemed legitimate from a list of nine. Six direct: 1) blame the chief, 2) threaten the chief, 3) refuse collective labour, 4) refuse to work on the chief's farm, 5) refuse marriage to chief's family, 6) refuse to pay local tax. Three

¹⁰ Comprehension questions were asked after outcome measures. Robustness checks excluding respondents who did not understand the experimental condition are in Appendix J.

¹² Unlike studies using open questions for priming, we asked this to half the sample across all conditions to validate the comprehensiveness of closed-ended responses in Outcomes 4 and 5.

indirect: 7) complain to an elder, 8) complain to section chief, 9) complain to paramount chief. Item order was randomized¹³¹⁴;

- Outcome 5 (main): Number of those sanctions respondents feel able to personally undertake.

The list of sanctions was developed from a literature review and field consultations with two groups (10 and 20 participants). The literature provided examples of citizen-led sanctions and identified sources of chiefs' authority, which we reframed as potential leverage points. Key informants reviewed and refined this list for local plausibility (See Appendix A for a formal definition of each sanction). The open-ended responses (Outcome 3), categorized in Appendix H, confirmed coverage of most relevant actions. Additional responses—e.g., “mobilize citizens” (3%), “report to police” (4%), “go to court” (2%), and “report to NGO” (2%)—were excluded due to low frequency.

We distinguish between perceived legitimacy (Outcome 4) and self-reported ability (Outcome 5) for two reasons. First, given the influence of chiefs in southern Sierra Leone, respondents may hesitate to express direct intentions; therefore, questions about legitimacy provide a less confrontational means to gauge preferences. This approach helps address concerns about social desirability bias. Second, little is known about non-electoral sanctions in traditional institutions; the legitimacy question maps the *repertoire* of politically acceptable responses. The ability question captures individual agency to act on these preferences. Appendix D details measurement strategy and summary statistics. To test hypotheses 1 and 2, we aggregate counts of direct and indirect sanctions.

Because the legitimacy and efficacy measures referenced that ‘some members of the village reacted in this way,’ it is possible that this phrasing introduced subtle demand effects by signaling socially acceptable responses. Nevertheless, it is unlikely to vary systematically between Treatment 1 and Treatment 4 where only the quantity of money taken changes. We develop sensitive analysis to demand effects in the analysis section.

¹³ The term legitimacy was translated using words meaning “right” or “fair” in Sherbro, Mende, and Krio.

¹⁴ We first asked the respondents whether each sanctions were felt as legitimate. Then we converted each individual answers to a number. The same holds for outcome 5.

Empirical strategy

We estimate the effect of providing information about leaders' malevolence on attitudes towards the legitimacy of sanctioning behavior with an average treatment effect estimand. As there is covariate balances between the control and the treatment group, we use the following estimator, for respondent j :

$$GSI_j = \beta_0 + \beta_1 Z_j + \gamma_g + \gamma_v + \epsilon_j \quad (1)$$

With GSI_j , the outcome variable, is the number of direct or indirect sanctions chosen by the respondent j , β_1 is the Average Treatment Effect, and Z_j is a dummy variable indicating whether the participant j belongs to treatment 1 or the control group, and to treatment 4 or the treatment 1 for the hypothesis 1 or the treatment arm 3 or 2 for the hypothesis 2. γ_g and γ_v are gender and village fixed effects accounting for the block randomization strategy¹⁵. We use robust HC2 standard errors (Aronow & Middleton 2013). The robustness checks include the pre-analysis plan strategy (excluding experimental block fixed effects), clustering standard errors at the village level, adding covariates that were not initially balanced, and excluding participants who did not fully understand the scenario (Appendix J). We also conducted p-value corrections for multiple hypothesis testing (in the main text) and performed a sensitivity analysis to account for potential experimental demand effects. Results shown in the main text for the hypotheses 1 and 2 are robust to all those alternative strategies. The experiment is well-powered when comparing treatment 1 with the control group, or treatment 4 with treatment 1. Power reaches 51% when comparing treatment 3 with treatment 2.

Results

The experimental conditions represent a clear breach of community norms. As Figure 1 shows, the proportion of respondents who disagree with the chief's behavior increases with the magni-

¹⁵ The pre-registered specification does not include experimental block fixed effects. Gerber & Green (2012) demonstrate that such a procedure increases precision, which is essential for testing hypothesis 2, where the power is around 50%. The pre-registered strategy is also reported in the Appendix and yields similar results.

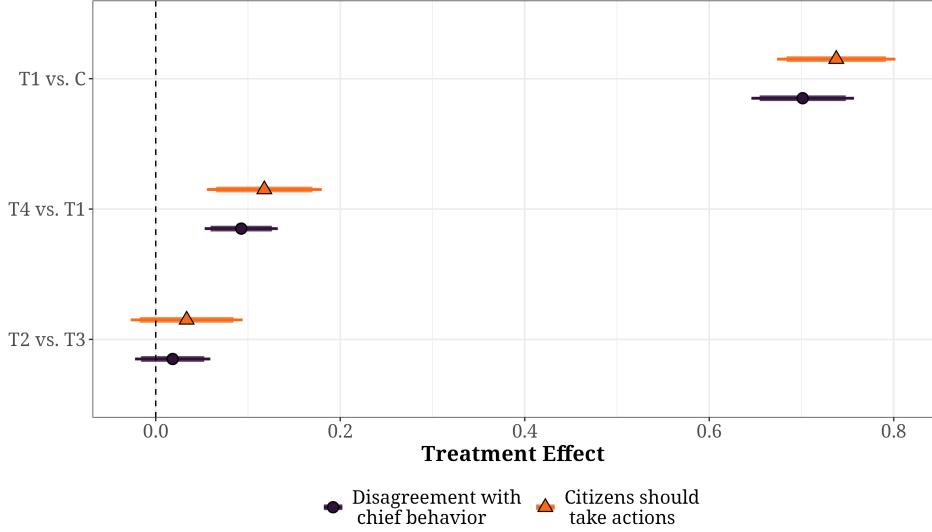


Figure 1: Estimated Treatment Effects: Disagreement with Chiefs' Behavior and Citizens' Actions

Notes: The figure presents estimated average treatment effects for two outcomes: (i) disagreement with the chiefs' behavior (a five-point Likert scale rescaled to range from 0 to 1) and (ii) whether citizens believe they should take any action (binary). The comparisons shown are Treatment 1 vs. Control, Treatment 2 vs. Treatment 3, and Treatment 4 vs. Treatment 1. Dots represent point estimates, thick bars indicate 95% confidence intervals, and thin bars indicate 90% confidence intervals. Treatment 1 corresponds to a scenario where chiefs steal a very small portion of development funds; Treatment 2 corresponds to a scenario where elders blame the chief after misappropriation; Treatment 3 corresponds to a scenario where elders do not blame the chief; and Treatment 4 corresponds to a scenario where half of the funds are stolen.

tude of the offense, reaching 70% in Treatment 1 ($\beta = 0.7, p < 0.001$) and 80% in Treatment 4 ($\beta = 0.1, p < 0.001$). However, this proportion does not vary with the elders' behavior ($\beta = 0.01, p = 0.377$). Furthermore, the proportion of respondents who believe that citizens should take action against the town chief also increases with the magnitude of the offense, reaching 74% in Treatment 1 ($\beta = 0.74, p < 0.001$) and 86% in Treatment 4 ($\beta = 0.12, p < 0.001$). This suggests that respondents want the norm to be enforced. Yet, this proportion likewise does not vary with the elders' behavior¹⁶ ($\beta = 0.03, p = 0.277$).

A wide range of sanctions is seen as legitimate

As shown by figure 2, respondents support sanctioning leaders via multiple channels and adjust their responses based on the magnitude of the offense, confirming hypothesis 1. Compar-

¹⁶ Appendix I displays formally the results for those outcomes accounting for multiple comparisons and multiple estimation strategies.

ing treatment 1 (minor theft) to the control reveals a significant increase of 3.04 ($\beta = 3.04$, $p < 0.001$) in the number of sanctions deemed legitimate (see Figure I.1 for breakdowns by sanction). Disaggregated effects show increases of 1.2 for direct sanctions and 1.9 for indirect sanctions. Comparing treatment 4 (major theft) to treatment 1, the number of legitimate sanctions rises by 0.54 ($\beta = 0.54$, $p < 0.001$), 0.27 for direct and 0.27 for indirect sanctions. Respondents are more likely to endorse a range of sanctions as legitimate when the theft is more severe. These findings remain robust when the alternative outcome—the number of feasible sanctions—is used¹⁷.

Figure I.1 in Appendix I further explores sanction preferences. The most commonly supported actions—endorsed by about 75% of respondents in treatment groups—involve appealing to higher authorities (elders, section chiefs, or the paramount chief). Blaming the chief directly or refusing to work on their farm are also frequently chosen but to a lesser extent. In contrast, the other direct sanctions—such as physical threats, tax refusal, or rejecting marriage ties—are rarely seen as legitimate, despite occasional references in the literature (Bulte et al. 2018; Richards 2021).

Appendix J.6 tests the hypotheses using the proportion of direct and indirect sanctions chosen. It reveals a statistically significantly larger effect size for the indirect-sanctions outcomes. Appendix K includes a formal statistical test that confirms the preference for indirect sanctions. This evidence is also supported by the open ended survey question. A random subset of respondents was asked how they would respond in such a situation, and all indicated an indirect approach: they would complain to a tribal authority, an elder, or the section/paramount chief. The low incidence of direct sanctions can be equally explained by interpersonal conflict avoidance and by villagers' high baseline satisfaction (94%) and trust in their town chiefs—who are often the most accessible and responsive authorities—rather than an unwillingness to act.

¹⁷ Appendix J displays formally the results for those outcomes accounting for multiple comparisons and multiple estimation strategies.

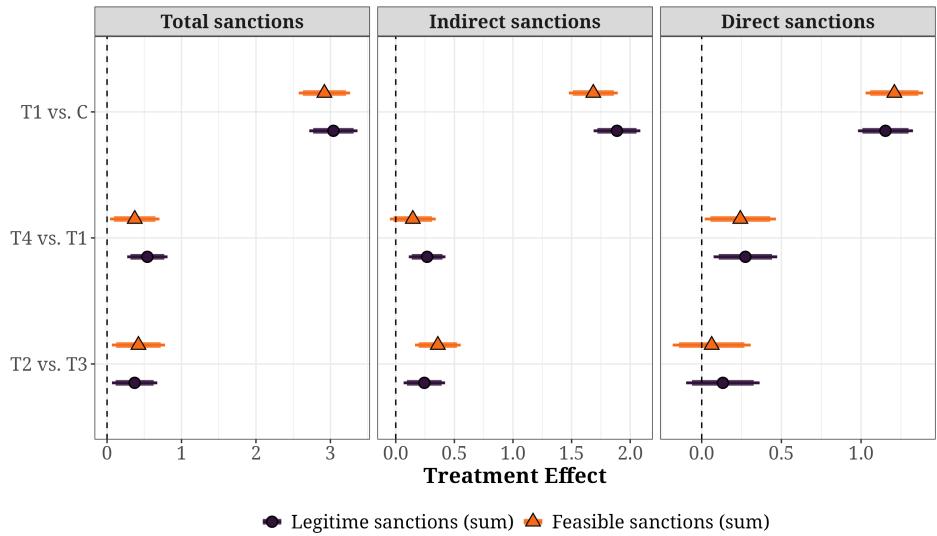


Figure 2: Estimated Treatment Effects on the Number of Legitimate and Feasible Sanctions

Notes: The figure presents estimated average treatment effects for two outcomes: (i) total number of legitimate sanctions, the number of direct and indirect legitimate sanctions against the village chief and (ii) the total number of feasible sanctions, the number of direct and indirect feasible sanctions against the village chief. The comparisons shown are Treatment 1 vs. Control, Treatment 2 vs. Treatment 3, and Treatment 4 vs. Treatment 1. Dots represent point estimates, thick bars indicate 95% confidence intervals, and thin bars indicate 90% confidence intervals. Treatment 1 corresponds to a scenario where chiefs steal a very small portion of development funds; Treatment 2 corresponds to a scenario where elders blame the chief after misappropriation; Treatment 3 corresponds to a scenario where elders do not blame the chief; and Treatment 4 corresponds to a scenario where half of the funds are stolen.

Elders' involvement increases perceived legitimacy

Contrary to expectations, respondents do not view fewer sanctions as legitimate when elders have already blamed the chief. Instead, they are more likely to regard a wider range of sanctions as legitimate in such cases. Comparing treatment 2 (where elders blamed the chief) and treatment 3 (where they did not), the number of legitimate sanctions increased by 0.37 ($\beta = 0.37$, $p = 0.028$), driven primarily by a 0.24 increase in indirect sanctions. Appendix J.6 robustly reveals a statistically significantly larger effect size for the indirect-sanctions outcomes. While direct sanctions also rose, the effect was not statistically significant. These findings are robust when restricted to villages with a council of elders (around 90% of the sample). In the treatment 2 (where elders blamed the chief) vs. treatment 1 (which made no mention of elder behavior) comparison, effects are similar: 0.38 for total sanctions ($\beta = 0.38$, $p = 0.028$), 0.24 for direct ($\beta = 0.24$, $p = 0.028$), and 0.14 for indirect ($\beta = 0.14$, $p = 0.13$).

We identify two mechanisms: the information-legitimation mechanism, where elders' actions against the chief legitimize villagers' grievances, and the covering mechanism, where collective actions reduce retaliation risk. Appendix M displays formal evidence from statistical tests and from open-ended responses and suggests both mechanisms are at play, with elders' influence playing a key role in shaping villagers' willingness to challenge the chief and maintain social order. These findings contribute to a growing body of research demonstrating that traditional elites can promote mobilization and cooperation by rewarding such behavior (Goist & Kern 2018), while also shaping public opinion (Yekple & Mitkov 2024) and influencing political behavior (Brierley & Ofosu 2023).

Legitimacy of sanctions differs by social status

As shown in Figure 3, treatment effects vary systematically across subgroups. Comparing Treatment 1 to the control, men display larger increases than women in both direct and indirect sanctions considered legitimate, confirming that gender moderates sanctioning responses. Income further shapes responses: higher-income respondents (Q3–Q4) report weaker increases in indirect and direct sanctions deemed legitimate than poorer respondents, with a difference of about 1.0 for both direct and indirect sanctions between the top and lowest quartiles. This may

suggest that wealthier individuals are more tolerant of small theft. Voting rights for the town chief are also a strong moderator: respondents with voting rights show significant increases in both direct and indirect sanctions considered legitimate, with large (about 1.0) and statistically significant differences between these two groups. By contrast, comparisons between Treatment 4 and Treatment 1, and between Treatment 2 and Treatment 3, yield no systematic heterogeneity across subgroups. These findings suggest that socioeconomic and political status influence how citizens view the legitimacy of sanctions against chiefs.

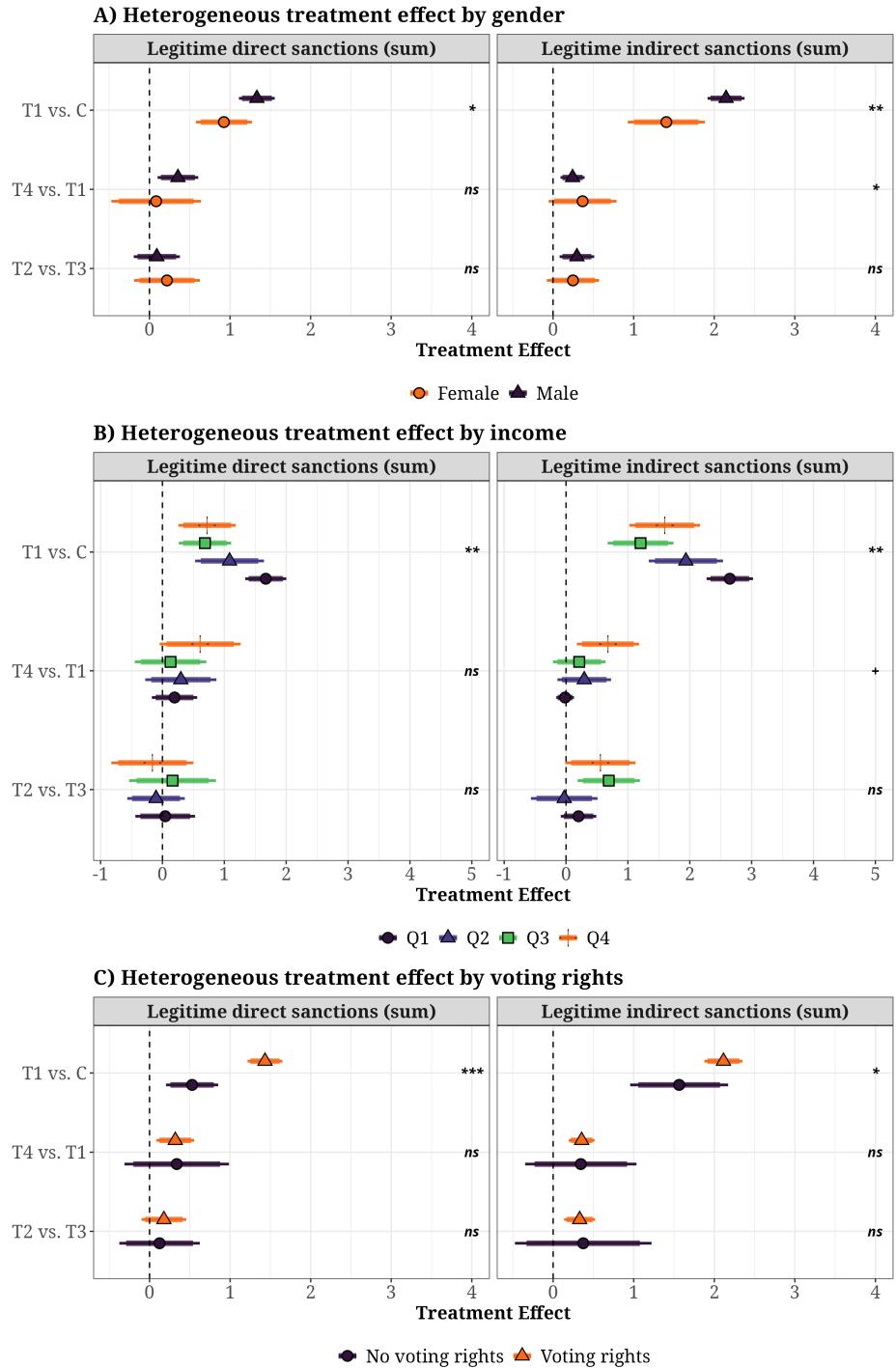


Figure 3: Heterogeneity analysis of the estimated treatment of effect on the number of legitimate sanctions by gender, income, and voting rights

Notes: Panel A presents the average values of total, direct, and indirect sanctions considered as legitimate in the control, treatment 1, and treatment 4 conditions, with the dots representing these averages. The bars illustrate two standard errors. In Panel B, the figure displays the average legitimate sanctions per sanction type and treatment condition. Since the outcome in this case is a binary variable, no standard errors are shown. Treatment 1 corresponds to the scenario where the chiefs steal a very small portion of development funds while treatment 4 represents a scenario where half of the money is stolen.

Our experimental design does not allow us to draw substantial conclusions from the heterogeneity analysis. However, if town chiefs are aware of and responsive to these potential sanctions, differences between subgroups may help explain why current local political choices tend to favor men and citizens with traditional political rights (insiders).

Limitations

We test the robustness of our results to experimenter demand effects, which occur when subjects infer the expectations of researchers and adjust their behavior accordingly (Mummolo & Peterson 2019). Our survey experiment was embedded within a one-hour household survey primarily focused on understanding livelihood activities and deforestation patterns in these communities. This design provides strong reason to believe that respondents were unlikely to discern the experimenter's intentions, thereby reducing the likelihood of such effects. Nonetheless, our empirical strategy addresses this concern through a bounding estimation strategy. Previous studies have found demand effects to be either negligible (Mummolo & Peterson 2019) or relatively small (de Quidt et al. 2018). We apply a bounding estimation strategy, accounting for these effects using the 0.1 to 0.3 standard deviation range reported by de Quidt et al. (2018). The results are displayed in Appendix J. Sensitivity analysis of hypothesis 1 demonstrates that the results are robust to all experimenter demand effect sizes, while for hypothesis 2, small experimenter demand effects could explain the findings. Furthermore, because our design does not include a condition where elders' behavior is described in the absence of chief misappropriation, we cannot fully disentangle the independent effect of elders' actions from the possibility that longer vignettes themselves cue participants to reflect more carefully.

Conclusion

This paper investigates how villagers in Sierra Leone perceive the legitimacy and feasibility of non-electoral sanctions against village chiefs. Using a survey experiment, we find that indirect sanctions—such as reporting to higher authorities—are viewed as more legitimate and feasible than direct actions. The range of sanctions considered acceptable increases with the severity

of the chief's offense, and preferences vary by gender, income, and voting rights: men, lower-income individuals, and those with voting rights endorse a broader set of sanctions. Elders also play a key role; when they publicly criticize a chief, villagers expand the set of sanctions they deem acceptable.

These findings suggest that even where electoral mechanisms are rare or absent, villagers perceive avenues to discipline chiefs, primarily mediated through hierarchical institutions rather than direct action. While consistent with accounts emphasizing chief accountability to higher authorities (Acemoglu et al. 2014; Bulte et al. 2018; Labonte 2012), our results show that community members retain some agency, leveraging hierarchical relationships or occasionally imposing direct sanctions such as refusing to farm on the chief's land. Whether such sanctioning norms are effective in promoting downward accountability remains an open question.

The study also contributes to debates on accountability under traditional authority (Baldwin 2025), highlighting that villagers recognize a wide range of legitimate sanctions and that elders function as political intermediaries, legitimizing grievances and reducing sanctioning risks (Baldwin et al. 2022).

These results should be interpreted cautiously, as they reflect responses to hypothetical misappropriation and may not generalize to other types of offenses. Endorsement of sanctions may not translate into implementation, which depends on power dynamics or coordination. Future research could examine links between perceived legitimacy and actual sanctioning, explore other governance systems, and investigate how gender, wealth, and voting rights shape power and inequality.

Data availability

Funding was received from the green transition scheme of the International Growth Center. The data, code, and any additional materials required to replicate all analyses in this article are available at the Journal of Experimental Political Science Dataverse within the Harvard Dataverse Network, at: doi: <https://doi.org/10.7910/DVN/K8QXXV> (Chazottes & Nabieu 2025).

Acknowledgement

We thank Alphious Jalloh, Emmanuel Saffa, Jeremiah J.P. Simbo, John Jusu, Kadiatu Kanneh, Margret Kawa, Mariama M Barrie, Mohamed Lawan, Samai Lahai, Sayo K. Mansaray, Sheku Bengura, and Tommy Barley for excellent research assistance. We are strongly thankful for the logistical support and feedback received from Niccolo Meriggi, and the Wageningen University teams in Sierra Leone. We have received valued feedback from Maarten Voors, Miriam Golden, Susanna Garside, Virginia Rocha, Daniel Goldstein, and the other attendees of the Political Behavior Colloquium at the European University Institute, and Michaël Aklin at EPFL Lausanne.

Conflicts of Interest.

The authors have no competing interests to declare.

Ethics Statement

The research protocol for this study received approval from the Office of Sierra Leone Ethics and Scientific Review Committee (SLESRC), under protocol number n°020/04/2023. A copy of the official approval letter has been uploaded to the editorial management system (EM).

This research was conducted in adherence to the American Political Science Association (APSA)'s Principles and Guidance for Human Subjects Research.

Enumerators have first read an Informed Consent form to the respondents and ensure that they had read the entire informed consent, understood it, and agreed before proceeding. Respondents were informed that their participation was completely voluntary and that they were allowed to withdraw from the study at any time without any impact on their relationship with the academic institution. Respondents were also informed that they were able to abstain from answering any question. This informed consent document also informed respondents that there were no known risks or direct benefits of the study. Respondents were assured that no identifying information about them would be made public and any views they expressed would be

kept completely confidential. A full copy of the informed consent is available by contacting the authors.

The studies did not use deception in this experiment. Nor did the survey experiment intervene in political processes. Respondents were given the field coordinator and the IRB's contact information for any concerns or questions. Respondents were not compensated for their participation in the study.

References

- Acemoglu, D., Reed, T., & Robinson, J. A. (2014). Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone. *Journal of Political Economy*, 122(2), 319–368. Publisher: The University of Chicago Press.
- Anderson, S. E., Buntaine, M. T., Liu, M., & Zhang, B. (2019). Non-Governmental Monitoring of Local Governments Increases Compliance with Central Mandates: A National-Scale Field Experiment in China. *American Journal of Political Science*, 63(3), 626–643. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12428>.
- Anti-Corruption Commission of Sierra Leone (2024). AAC calls on Tankoro Chiefdom Development Committee to maintain transparency and accountability.
- Arnall, A., Thomas, D. S. G., Twyman, C., & Liverman, D. (2013). NGOs, elite capture and community-driven development: perspectives in rural Mozambique*. *The Journal of Modern African Studies*, 51(2), 305–330. Publisher: Cambridge University Press.
- Aronow, P. M., & Middleton, J. A. (2013). A Class of Unbiased Estimators of the Average Treatment Effect in Randomized Experiments. *Journal of Causal Inference*, 1(1), 135–154. Publisher: De Gruyter Section: Journal of Causal Inference.
- Azfar, O., & Gurgur, T. (2008). Does corruption affect health outcomes in the Philippines? *Economics of Governance*, 9(3), 197–244.
- Balasuriya, K. (2023). Engaging Traditional Leaders in Local Governance: Devising a Generalisable Framework for Sub-Saharan Africa. *The Journal of Development Studies*, 59(5), 637–652. Publisher: Routledge _eprint: <https://doi.org/10.1080/00220388.2023.2172332>.
- Baldwin, K. (2016). *The Paradox of Traditional Chiefs in Democratic Africa*. Cambridge University Press. Google-Books-ID: u23OCgAAQBAJ.
- Baldwin, K. (2025). Traditional Authorities Around the World: Toward a Better Understanding of Their De Facto Power, Downward Accountability, and State Recognition. *Annual Review of Political Science*, 28(Volume 28, 2025), 95–114. Publisher: Annual Reviews.
- Baldwin, K., & Holzinger, K. (2019). Traditional Political Institutions and Democracy: Reassessing Their Compatibility and Accountability. *Comparative Political Studies*, 52(12), 1747–1774. Publisher: SAGE Publications Inc.
- Baldwin, K., Muyengwa, S., & Mvukiyehe, E. (2022). Accountability and Inclusion in Customary Institutions: Evidence from a Village-Level Experiment in Zimbabwe. *Quarterly Journal of Political Science*, 17(2), 283–315. Publisher: Now Publishers, Inc.
- Baldwin, K., & Ricart-Huguet, J. (2022). Does Land Quality Increase the Power of Traditional Leaders in Contemporary Africa? *The Journal of Politics*.

- Becorpi, C. (2018). Enhancing participatory conflict transformation in Sierra Leone: are ‘traditional’ authorities the best game in town? *African Studies*, 77(1), 105–126. Publisher: Routledge eprint: <https://doi.org/10.1080/00020184.2017.1420337>.
- Brierley, S., & Ofosu, G. K. (2023). Chiefs’ Endorsements and Voter Behavior. *Comparative Political Studies*, (p. 00104140231194916). Publisher: SAGE Publications Inc.
- Bulte, E., Richards, P., & Voors, M. (2018). *Institutions and Agrarian Development: A New Approach to West Africa*. Springer. Google-Books-ID: rk1yDwAAQBAJ.
- Carlson, E. (2021). Distribution or Diversion? Distribution of Local Public Goods in the Presence of Clientelist Brokers. *Political Research Quarterly*, 74(2), 274–287. Publisher: SAGE Publications Inc.
- Carlson, E., & Seim, B. (2020). Honor among Chiefs: An Experiment on Monitoring and Diversion Among Traditional Leaders in Malawi. *The Journal of Development Studies*. Publisher: Routledge.
- Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*. *The Quarterly Journal of Economics*, 127(4), 1755–1812.
- Chazottes, R., & Nabieu, J. (2025). Replication Data for: Non-Electoral Accountability: Citizen Sanctions on Traditional Leaders in Sierra Leone. DOI: 10.7910/DVN/K8QXXV.
- Chen, J., Pan, J., & Xu, Y. (2016). Sources of Authoritarian Responsiveness: A Field Experiment in China. *American Journal of Political Science*, 60(2), 383–400. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12207>.
- Conteh, M. N. (2013). The Institution of Paramount Chieftaincy in Sierra Leone. An Introduction to its History and Electoral Process. Tech. rep., National Electoral Commission, State Avenue, OAU Drive, Freetown.
- de Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and Bounding Experimenter Demand. *American Economic Review*, 108(11), 3266–3302.
- Fantherope, R. (2006). On the limits of liberal peace: Chiefs and democratic decentralization in post-war Sierra Leone. *African Affairs*, 105(418), 27–49.
- Gerber, A. S., & Green, D. P. (2012). *Field Experiments: Design, Analysis and Interpretation*. New York: Norton.
- Goist, M., & Kern, F. G. (2018). Traditional institutions and social cooperation: Experimental evidence from the Buganda Kingdom. *Research & Politics*, 5(1), 2053168017753925. Publisher: SAGE Publications Ltd.
- Grant, R. W., & Keohane, R. O. (2005). Accountability and Abuses of Power in World Politics. *American Political Science Review*, 99(1), 29–43.
- Grieco, K. (2024). Building Fiscal Capacity with Traditional Political Institutions: Experimental and Qualitative Evidence from Sierra Leone. report, The Institute of Development Studies and Partner Organisations.
- Henn, S. J. (2022). Complements or Substitutes? How Institutional Arrangements Bind Traditional Authorities and the State in Africa. *American Political Science Review*, (pp. 1–20). Publisher: Cambridge University Press.
- Honig, L. (2017). Selecting the State or Choosing the Chief? The Political Determinants of Smallholder Land Titling. *World Development*, 100, 94–107.

- Honig, L. (2022). *Land Politics: How Customary Institutions Shape State Building in Zambia and Senegal*. Cambridge University Press.
- Humphreys, M., & Weinstein, J. M. (2006). Handling and Manhandling Civilians in Civil War. *American Political Science Review*, 100(3), 429–447. Publisher: Cambridge University Press.
- Labonte, M. T. (2012). From patronage to peacebuilding? Elite capture and governance from below in Sierra Leone. *African Affairs*, 111(442), 90–115.
- Leach, M. (2022). Rainforest Relations: Gender & Resource Use by the Mende of Gola, Sierra Leone. In *Rainforest Relations*. Edinburgh University Press.
- Mattingly, D. C. (2016). Elite Capture: How Decentralization and Informal Institutions Weaken Property Rights in China. *World Politics*, 68(3), 383–412. Publisher: Cambridge University Press.
- Mbiti, I. M. (2016). The Need for Accountability in Education in Developing Countries. *Journal of Economic Perspectives*, 30(3), 109–132.
- Meng, A., Paine, J., & Powell, R. (2023). Authoritarian Power Sharing: Concepts, Mechanisms, and Strategies. *Annual Review of Political Science*, 26(Volume 26, 2023), 153–173. Publisher: Annual Reviews.
- Mesquita, B. B. D., Smith, A., Siverson, R. M., & Morrow, J. D. (2005). *The Logic of Political Survival*. MIT Press. Google-Books-ID: HBiv9Iy9zYIC.
- Mihaylova, I. (2023). Perpetuating the malign legacy of colonialism? Traditional chiefs' power and deforestation in Sierra Leone. *World Development*, 164, 106176.
- Mokuwa, E., Voors, M., Bulte, E., & Richards, P. (2011). Peasant grievance and insurgency in Sierra Leone: Judicial serfdom as a driver of conflict. *African Affairs*, 110(440), 339–366.
- Mummolo, J., & Peterson, E. (2019). Demand Effects in Survey Experiments: An Empirical Assessment. *American Political Science Review*, 113(2), 517–529.
- Murtazashvili, I., & Murtazashvili, J. (2016). The origins of private property rights: states or customary organizations? *Journal of Institutional Economics*, 12(1), 105–128. Publisher: Cambridge University Press.
- Njong, A. M., & Ngantcha, J. T. (2013). Institutions and Leakage of Public Funds in the Cameroonian Healthcare Delivery Chain. *Journal of African Development*, 15(1), 19–43. Publisher: Duke University Press.
- Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives*, 14(3), 137–158.
- Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton University Press. Publication Title: Understanding Institutional Diversity.
- Paller, J. W. (2014). Informal Institutions and Personal Rule in Urban Ghana. *African Studies Review*, 57(3), 123–142.
- Peters, K. (2011). The Crisis of Youth in Postwar Sierra Leone: Problem Solved? *Africa Today*, 58(2), 129–153. Publisher: Indiana University Press.
- Reinikka, R., & Svensson, J. (2002). Explaining Leakage of Public Funds.

- Richards, P. (2005). To fight or to farm? Agrarian dimensions of the Mano River conflicts (Liberia and Sierra Leone). *African Affairs*, 104(417), 571–590.
- Richards, P. (2021). Public authority and its demons: the Sherbro leopard murders in Sierra Leone. *Africa*, 91(2), 226–248. Publisher: Cambridge University Press.
- Ruppel, S., & Leib, J. (2022). Same but different: the role of local leaders in the peace processes in Liberia and Sierra Leone. *Peacebuilding*, 10(4), 470–505. Publisher: Routledge eprint: <https://doi.org/10.1080/21647259.2022.2027152>.
- Sawyer, E. (2008). Remove or Reform? a Case for (Restructuring) Chiefdom Governance in Post-Conflict Sierra Leone. *African Affairs*, 107(428), 387–403.
- Scott, J. C. (1985). *Weapons of the Weak: Everyday Forms of Peasant Resistance*. Yale University Press. Google-Books-ID: wRcOEAAAQBAJ.
- Tsai, L. L. (2007). *Accountability without Democracy: Solidary Groups and Public Goods Provision in Rural China*. Cambridge University Press. Google-Books-ID: B0oxnuDMI2EC.
- Voors, M., Turley, T., Bulte, E., Kontoleon, A., & List, J. A. (2018). Chief for a Day: Elite Capture and Management Performance in a Field Experiment in Sierra Leone. *Management Science*, 64(12), 5855–5876. Publisher: INFORMS.
- Wilfahrt, M. (2018). The politics of local government performance: Elite cohesion and cross-village constraints in decentralized Senegal. *World Development*, 103, 149–161.
- Yekple, M., & Mitkov, Z. (2024). Can trusted authorities change minds on anti-LGBTQ norms: Evidence from an experiment in Ghana. *PLOS ONE*, 19(8), e0304698. Publisher: Public Library of Science.
- Zhong, P., & Zeng, Y. (2024). Does top-down accountability promote responsiveness? Evidence from a survey experiment in China. *Journal of Chinese Governance*, 9(1), 1–22. Publisher: Routledge eprint: <https://doi.org/10.1080/23812346.2023.2300181>.

Supplementary materials for "Non-Electoral Accountability: Citizen Sanctions on Traditional Leaders in Sierra Leone"

Rens Chazottes¹, Junisa Nabieu

October 2025

Table of Contents

A Theoretical framework	4
B Institutional context	8
B.1 Chieftaincy organization	8
B.2 Descriptive statistics	9
C Pre-Analysis Plan	10
C.1 Theory and hypotheses	11
C.2 Theoretical expectations	12
C.3 Empirical strategy	14
C.4 Measurement strategy	17
C.5 Analysis	19
D The study area, measurement strategy and summary statistics	22
D.1 Study area	22
D.2 Dependent variables used in hypothesis 1 and 2	23
D.3 Descriptive statistics of the respondents	24

¹ rens.chazottes@proton.me

D.4	Inequality between town chiefs and community members	25
E	Covariate balances	26
F	Understanding of experimental scenarios	27
G	Deviation from the pre-analysis plan	28
G.1	Sampling and experimental conditions	28
G.2	Inclusion and wording of the hypotheses and deviations	29
G.3	Measurement analysis and deviations	30
G.4	Empirical strategy and deviations	31
G.5	Pre-Registered and Main Text Results Overview	31
G.6	Main results using the land deal survey experiment	32
H	Open-question outcome summary	35
I	Treatment effects on the main outcomes	37
I.1	Outcome 1: disagreement with the chief's behavior	37
I.2	Outcome 2: Citizens should take actions against the chief	39
I.3	Legitimacy of the sanctions	41
I.4	Feasibility of the sanctions	44
J	Robustness check for the statistical test of hypotheses 1 and 2	47
J.1	Comparison of treatment 1 with the control arm	47
J.2	Comparison of treatment 4 with the treatment 1	48
J.3	Sensitivity analysis to experimenter demand effect (EDE) - hypothesis 1	49
J.4	Robustness check for the statistical test of hypotheses 2	50
J.5	Sensitivity analysis to experimenter demand effect (EDE) - hypothesis 2	51
J.6	Robustness check with alternative measurement strategy: proportion of sanctions chosen	51
K	Relative costs and legitimacy of sanctions	53
K.1	Summary statistics of the outcomes of the focus group discussions	53
K.2	Relationships between legitimacy of the sanctions and perceived costs	54
K.3	Statistical test of hypothesis 1a	55
K.4	Additional results	56
L	Heterogeneous effect by gender, income, and voting rights	57

L.1	Gender	57
L.2	Income	58
L.3	Voting rights	59
M	Elder involvement: information and covering mechanisms	60
N	Checklist for Reporting Standards for Experiments	62
N.1	Hypotheses	62
N.2	Subjects and Context	62
N.3	Allocation Method	63
N.4	Treatments	64
N.5	Results	65
N.6	Other Information	66

A Theoretical framework

Figure A.1 presents the conceptual model of non-electoral accountability in traditional chieftaincy. The framework begins with the norms of correct behavior, which establish the duties and expectations of chiefs within the community. A breach of these norms—such as the misuse of community resources—produces the first critical condition, namely that citizens recognize an “off-equilibrium” event has occurred. We measure this critical in our experimental setting using the following survey question: agreement with the chief’s behaviour using a 5-points Likert scale (outcome 1 in figure A.1).

From this breach, two conditions shape whether accountability mechanisms are activated. First, the availability of information determines whether citizens can reliably detect norm violations. This is not the focus of our empirical study. Second, citizens’ willingness to mobilize captures the degree to which community members are prepared to act on this information, despite potential costs or risks. We measure this condition using the following survey question: Whether citizens should respond (binary outcome, yes=1) (outcome 2 in figure A.1). Together, these two conditions shape the norms regarding correct sanctions, which specify socially recognized responses to violations.

These sanction norms are filtered through two additional dimensions. The legitimacy of sanctions (Outcome 4) reflects whether a given sanction is perceived as rightful and socially acceptable, while the feasibility of sanctions (Outcome 5) captures the practical capacity to implement them given material constraints, coordination problems, or fear of retaliation.

Conditional on legitimacy and feasibility, sanctions take one of two enforcement pathways. Direct sanctions are imposed by citizens themselves (e.g., withholding labor, refusing tax payments), while indirect sanctions rely on intermediaries such as elders, section chiefs, or paramount chiefs to discipline the offending leader. These enforcement pathways ultimately shape the chief’s adaptive behavior, as leaders adjust their conduct in anticipation of, or in response to, sanctioning pressures.

We believe that the identity of the enforcer is critical as the chief may be more likely to adjust its behaviour and be responsive to the identity of the expected enforcer of the norms instead of the community as a whole.

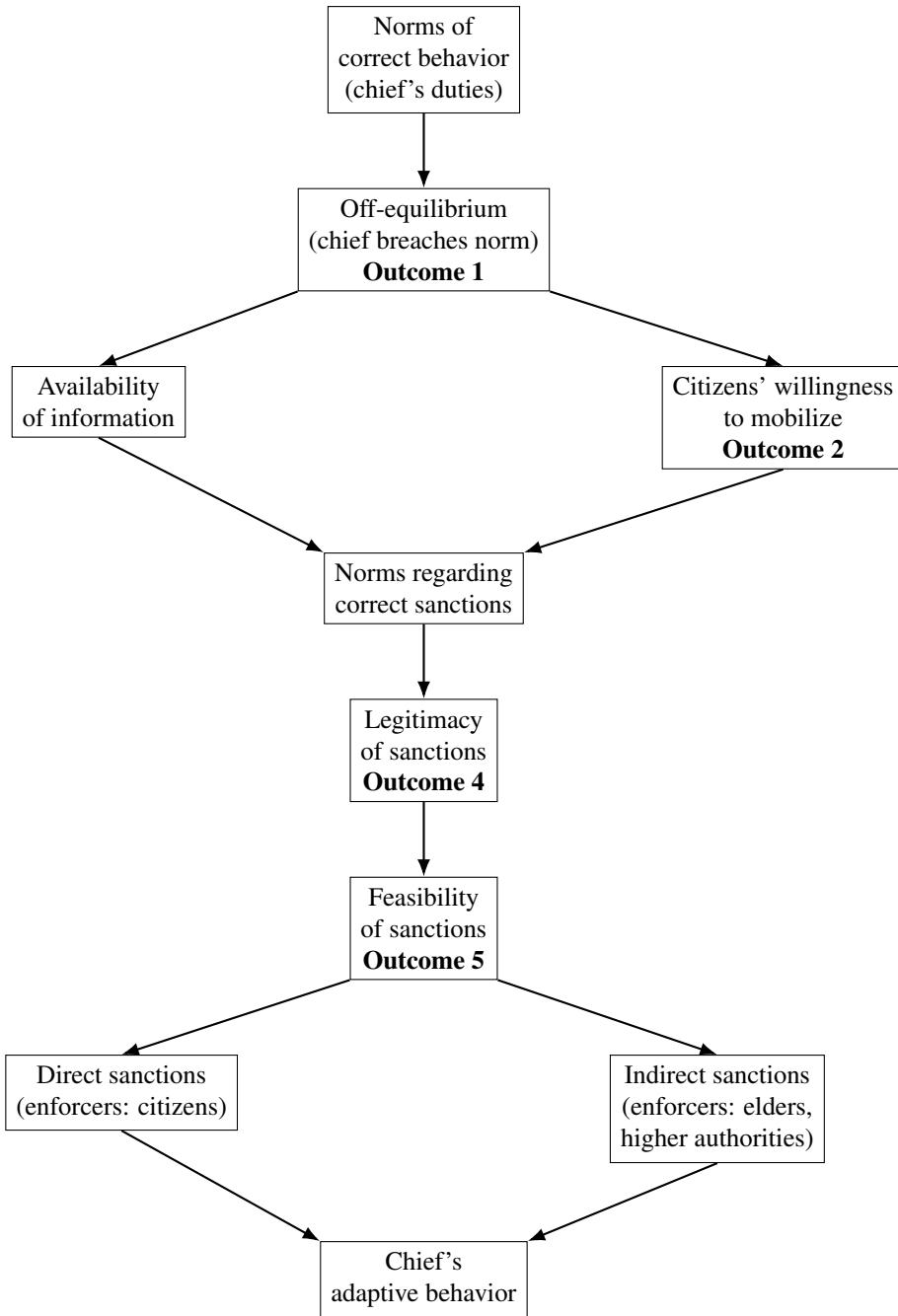


Figure A.1: Conceptual model of non-electoral accountability in traditional chieftaincy.

Note: Chiefs are expected to follow behavioral norms. When a chief breaches these norms and it is perceived as such by citizens (Outcome 1), two key conditions for non-electoral accountability need to be activated: the availability of information about the breach and citizens' willingness to mobilize (Outcome 2). These factors jointly influence the norms regarding correct sanctions. The perceived legitimacy of sanctions (Outcome 4) determines which sanctions are socially accepted, while feasibility (Outcome 5) constrains which sanctions can actually be implemented. Depending on these conditions, citizens may enact direct sanctions, and elders or higher authorities may enact indirect sanctions. These enforcement actions ultimately shape the chief's adaptive behavior, reflecting the dynamic feedback between norms, sanctioning practices, and observed breaches. Arrows indicate sequential processes.

We describe the *repertoire of political sanctions* available to community members as follow:

- **Blame the chief:** Publicly attributing fault to the chief for a perceived violation of community interests, often through verbal criticism in community meetings or informal gatherings. Blaming is symbolic but socially costly, as it undermines the chief's reputation and moral standing. Labonte (2012) acknowledges the existence of challenging chiefs through public forum but argues that is unlikely to happen in the Sierra Leonian context: "non-elites are risk-averse in claiming their rights from elites, airing grievances in public forums, or demanding accountability". Furthermore, Voors et al. (2018) mention that often such grievances may arise in private meeting inside secret society meetings.
- **Threaten the chief physically:** Issuing physical threats that more severe actions may follow if the chief does not change behavior. Conteh (2013) mentions the use of "revolts" and "murders". It also echoes instances of leopard murders in Sierra Leone as mentioned by Richards (2021): "Cannibalism is recognized from its physical manifestations – evidence of a body being cut open, often with scratches resembling the wounds caused by leopards, and the removal of entrails. Furthermore, there is a difference in the way the arrow of blame points (Douglas, 1992: 93). In witchcraft cases it is directed horizontally at peers; in cannibalism it points upwards, towards chiefs and elders". It also echoes other social accountability dimensions in other contexts, such as in Cameroon where it is widely acknowledge that physical threat and murder is a way to change leader (Claessen 2011).
- **Complain to an elder:** Bringing grievances about the chief's conduct to respected senior figures within the village (often lineage elders or village level advisory council members). Elders are custodians of custom and act as mediators between chiefs and villagers and act as an authority challenging the authority of the chief (Voors et al. 2018). Conteh (2013) further emphasizes that "consultations with the council of elders by a chief before taking major decisions" is a major dimension of check and balances within the chieftaincy system, and that every citizens as the right to complain to a figure with equal status of the chief to confront the town chief.
- **Complain to the section chief:** Escalating grievances to the section chief, who oversees multiple villages within a chiefdom section. Section chiefs exercise intermediate authority and are positioned to disanction village chiefs, pressure them to change behavior, or remove them from office. As Labonte (2012) mention: "a majority of taxpayers can remove a Town or Village Chief through their section chief and the Chiefdom Committee".
- **Complain to the paramount chief:** Appealing directly to the paramount chief, the highest customary authority in the chiefdom. paramount chiefs have formal jurisdiction to discipline or depose village chiefs, though their responsiveness varies. As Voors et al. (2018) mention: "If a majority of the taxpaying population is unsatisfied with their chief, they can try to remove him through higher-up layers in the traditional leadership system", referring to the work done by Labonte (2012).

- **Refuse collective labor:** Withdrawing participation from collective community work (e.g., road maintenance, farm clearing, or building construction) organized under the chief's authority. This is a powerful sanction because chiefs rely on collective labor to maintain legitimacy and deliver public goods. It echoes the following anecdotes from Fanthorpe (2006): "It was claimed in several other consultations that the people of particular chiefdom sections had adopted a policy of non-cooperation with their paramount or regent chief in protest against alleged misallocation of aid benefits". We make an argument that similar strategy may happen targetting village chief.
- **Refuse to work on the chief's farm:** Declining to provide labor on the chief's personal farm, a customary obligation that symbolizes respect and recognition of chiefly status. Refusal constitutes a direct challenge to the chief's authority and disrupts a key channel of personal benefit. Furthermore, in many places farmers rely on labour group for agriculture work. Denying access to labour can harm the socio-economic benefits of town chiefs. In Sierra Leone, Richards (2005) note that "abuses such as the seizure of the property of young widows, or the 'buying' of a case in order to exploit a young man's labour, should be challenged through exemplary appeals to higher courts".
- **Refuse marriage to chief's family:** Denying marriage alliances with the chief's family, thereby withholding important social and political ties that reinforce the chief's authority. Marriage networks are central to building legitimacy and alliances in Sierra Leone's kinship-based political order. As discussed by Richards (2021), "in the account offered by Bulte et al. (2018), freeborn families of the village republics were bonded by marriage exchanges between exogamous patriclans. This created a lifelong web of mutual obligations, to help with food security, in sickness, and eventually at death."
- **Refuse to pay local tax:** Withholding the annual local tax (levied through customary authority structures) that constitutes both a financial resource and a symbolic recognition of chiefly legitimacy. Tax refusal undermines the fiscal and moral basis of chiefly rule. Anecdotal evidence has shown such strategy being used in Sierra Leone and Mozambique (Arnall et al. 2013; Bulte et al. 2018; Fanthorpe 2006).

B Institutional context

B.1 Chieftaincy organization

The hierarchical structure of traditional authority extends beyond the village level, since each chiefdom is headed by a Paramount Chief who is supported by a Chiefdom Speaker serving as a deputy, and further comprises Section Chiefs and Town or Village Chiefs who preside over progressively smaller administrative units.

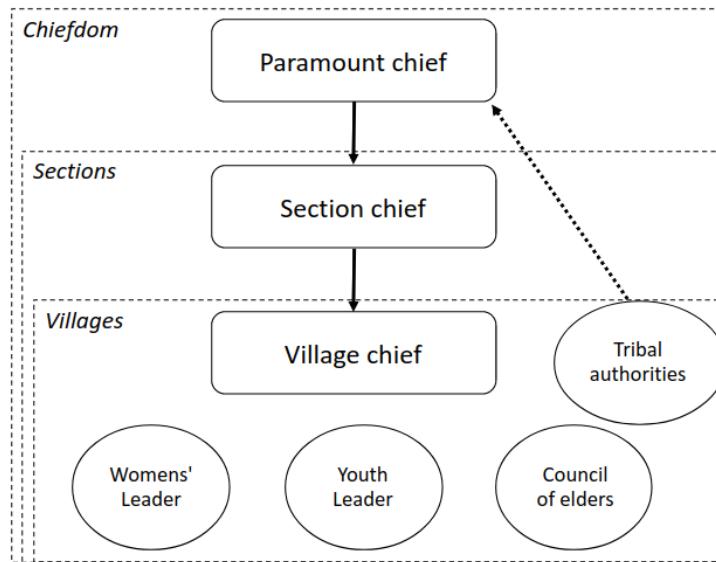


Figure B.1: Diagram of the chieftaincy structure in Sierra Leone.

B.2 Descriptive statistics

Table B.1: Chieftaincy characteristics in the Southern district and sample institutional and town chiefs' political characteristics

	N	mean	median	sd	min	max
Panel A: Attitudes towards chieftaincies (Afrobarometer data, round 8)						
Trust somewhat or a lot traditional leaders (Yes = 1)	240	0.54	-	-	0	1
Traditional leaders in competition with elected leader (Yes = 1)	240	0.31	-	-	0	1
<i>Power: At least some influence in</i>						
Governing your local community (Yes = 1)	240	0.88	-	-	0	1
Allocating land (Yes = 1)	240	0.76	-	-	0	1
Deciding your vote (Yes = 1)	240	0.32	-	-	0	1
Solving disputes (Yes = 1)	240	0.93	-	-	0	1
Panel B: Village level institutional characteristics (Casey et al. 2012)						
Town chief elected (Yes = 1)	80	0.47	-	-	0	1
Town chief can fine any offense (Yes = 1)	80	0.26	-	-	0	1
Fine for refusing supplying community labour (Yes = 1)	80	0.42	-	-	0	1
Fine for youth defiance (Yes = 1)	80	0.42	-	-	0	1
Panel C: Household level attitudes towards village chiefs (Casey et al. 2012)						
Satisfied with the town chief (Yes = 1)	959	0.94	-	-	0	1
Town chief fine people too much (Yes = 1)	959	0.14	-	-	0	1
Town chief powerful in development goods distribution (Yes = 1)	959	0.13	-	-	0	1
<i>Money related conflicts are brought to</i>						
The town chief court (Yes = 1)	959	0.47	-	-	0	1
The village elders (Yes = 1)	959	0.24	-	-	0	1
The local court, the magistrate court or the police (Yes = 1)	959	0.27	-	-	0	1
Panel D: Sample institutional characteristics						
Council of elders (Yes = 1)	77	0.87	-	-	0	1
Youth leader (Yes = 1)	77	1	-	-	0	1
Women leader (Yes = 1)	77	1	-	-	0	1
Previous town chief was suspended (Yes = 1)	77	0.24	-	-	-	-
Number of paramount chief ruling families	77	2.68	3	0.68	2	5
Number of village landowning families	77	4.37	3	3.54	0	17
Number of families having right to vote	77	7.89	6	5.62	1	30
Proportion of families having right to vote	77	0.85	1	0.23	0.12	1
Date of chief election	77	2012	2015	10	1982	2023
Succession period (in months)	77	11.05	8	14.11	0	84

Notes:

Panel A relies on round 8's Afrobarometer data using respondents from the Southern province. We provide summary statistics of the variables 41K, 87, and 90.

Panel B and C uses the data collection undertaken by Casey et al. Casey et al. (2012). Panel B focuses on the focus group discussions undertaken at the village level with village elites. Panel C summarizes variables collected at the household level.

Panel D comes from our own data collection in the sample of selected villages. It presents descriptive statistics of key institutional variables. The leadership position at the village level indicates the proportion of villages that have such a position. For instance, 86.84% of villages in the sample have a council of elders. We also provide the proportion of villages where a chief was suspended. It's important to note that families do not share the same political status in the study area. For each chiefdom (12 in total), we report the number of families eligible to compete for the role of paramount chief (paramount chief ruling families). At the village level (79 in total), we present the number of families that own land or can vote for the town chief, along with their respective percentages. Additionally, we include information on the date of the town chief election and the duration of the succession period.

C Pre-Analysis Plan

Rural Africa has the specificity of the co-existence of State and chieftaincies institutions in governing common affairs. Complementarities and competitions mark the development of these institutions. Debates remain vivid about the role of traditional leaders (or chiefs) in providing public goods, primarily because of the lack of accountability in traditional systems. However, the recent development in the chieftaincy literature underlines that traditional institutions have various features that are likely, in certain circumstances, to hold chiefs accountable and solve collective action dilemmas. Authors have focused on leaders' selection, the decision-making process, and their embeddedness in their communities (Baldwin 2016; Baldwin & Holzinger 2019). The classic democratic theory underlines the importance of competition for leaders' selection (Przeworski et al. 1999). In Sierra Leone, paramount chiefs facing more competitive elections are in chiefdoms with higher public good provisions (Acemoglu et al. 2014). Decision-making based on consensus, transparency, and inclusivity was underlined as essential for good leadership and the effectiveness of traditional political institutions (Börzel & Risse 2021; Magaloni et al. 2019). When well embedded in their communities, chiefs behave in close interdependent relationships with their community members, especially when they have long-time horizon (Dionne 2011). Social spaces sustained by reciprocity enable cooperation between group members through different channels. Such social space increases the expectation that peers will cooperate (Atwell & Nathan 2022) and sustain monitoring and trust without the need for sanctioning. However, even if community members are under scrutiny, such social spaces also enable a range of social sanctions (Miguel & Gugerty 2005; Tsai 2007).

All these features give rise to potential channels of bottom-up accountability mechanisms to enforce community norms. Bottom-up accountability mechanisms refer to citizens' actions to hold their leaders accountable. I mention these mechanisms as a strong sanctioning institution when these mechanisms are present. However, such accountability channels have not been the main focus of previous research on traditional leadership, cooperation, and public good provision (Baldwin & Holzinger 2019).

This research fills that gap by determining how citizens sanction undemocratic traditional village leaders and its consequences on town leaders' quality and resource governance effectiveness. The research is undertaken in small communities where chiefs are well embedded in their communities. It follows up a pre-registered pilot study undertaken in the Loma Mountain National Park in Sierra Leone. A household-level survey experiment and a village-level survey will be implemented in 80 villages in the Sherbro River Estuary in Sierra Leone. In two different situations, land allocation and community fund management, the experimental design disentangles how sanctioning mechanisms operate and grasps the diversity of sanctioning practices according to citizens' social and economic status. In a second step, the project explores how village-level heterogeneity in citizens' ability/willingness to sanction their village chief

impacts leaders' behavior and resource governance effectiveness measured as a deforestation rates index in the village's proximity.

C.1 Theory and hypotheses

Do sanctioning institutions shape leaders' behavior in acting for the common good? The early contributions about chieftaincies were clear regarding that matter. Because of upward accountability mechanisms and their colonial legacy, they conceptualized chiefs as colonial inventions behaving as despots (Mamdani 1996). They emphasized their low popular support (Ribot 2002) sustained by a lack of institutional alternatives for the population (Ntsebeza 2005). Logan Logan (2013) showed that empirical findings do not support this view. Indeed, using Afrobarometer data from 19 countries, she finds that traditional leaders "enjoy widespread popular legitimacy, and most believe that traditional authorities have an important role to play in local governance (Logan, 2013, p353)".

Recent works acknowledge that chiefs can behave as development brokers with a positive impact on their community Baldwin (2016). Along with other arguments, bottom-up accountability channels were discussed as one feature explaining the development broker model (Baldwin & Holzinger 2019). Such an argument used the classic conceptual framework of democratic theories, not without any problem (Neupert-Wentz et al. 2022). This conceptual stretching (Sartori 1970) might give a truncated vision on what basis chiefs are accountable to their community members. When fair and competitive, election is a tool to select representative leaders and punish those with bad records or poor outcomes (Przeworski et al. 1999). But in rural Africa, many traditional leaders are selected and rule for life. Being removed from office is an exception, not the rule. Instead, bottom-up sanctions would either pressure chiefs higher in the hierarchy to formally sanction the town chief (activation of upward accountability channels) or sanction directly the town chief through complex socio-economic institutions. The presence of reciprocity norms in communities where chiefs are well embedded give birth to this latter type of sanctioning institutions. This is the main argument that this experimental research tests

Reciprocity and sanctions

In this section, I lay out the various sanctions used against village chief that did not behave in the community interest. With such behavior, the village chief breaks reciprocity institutions. As a consequence community-members stop fulfilling their duties. The first sanction targets the ability to enjoy community-level reciprocity institutions. The communities rely primarily on farming and fishing activities that require collective organization. For example, most of the farms will rely on rotational labor groups for the harvesting of rice. Someone will work on someone else's farm because they expect the others to do the same in return. When people lose trust in someone, they close the institution to that person (Bulte et al. 2018).

The second sanction uses marriage institutions. Creating blood links with other families is a means of attracting resources and influence. When a chief loses respect, citizens will refuse to get married to the chief family. Doing so, they preserve their family from potential scandals and reduce the chief's overall influence.

The third set of sanctions targets, more specifically, the chief. The chief is a figure of authority within the village. When the chief is respected, it has the legitimacy to fulfill its duty. It can organize collective labor for village purposes, collect taxes, resolve conflicts, and implement and enforce by-laws. Questioning the chief's authority and not following its order is a strong message and has a significant negative impact on its power. Therefore, blaming the chief publicly, refusing to pay local taxes, or working for collective labor will fall under this type of sanctions (Arnall et al. 2013; Bulte et al. 2018).

Chiefs' removal from office

At the village level, the traditional structure comprises the town chief, tribal authorities (that can vote for the position of the paramount chief), the mummy queen, the youth leader, and the council of elders. The latter has an essential role in advising the chief in making decisions. Pressuring the council of elders can be an effective means for ordinary citizens to pressure the town chief. The chiefs can also be formally sanctioned within the chieftaincy structure. The paramount chief rules on the section chief, who rules on town chiefs. When there are issues between the village chief and the community, the section chief is the first authority called to resolve the dispute. If the dispute remains unresolved, the paramount chief is called. If the issue is severe, the town chief can be suspended for some time or removed from office. Consequently, when a chief has committed a theft, citizens inform authorities that can formally sanction the chief (Baldwin 2016). However, such actions hardly happen. Finally, as the chief rules for life, the only way to remove him from office in certain societies is to commit murder. As a result, in many instances in Cameroon, rebellions occurred, and chiefs were physically threatened or killed.

C.2 Theoretical expectations

C.2.1 Expectations

Consistent with evidence found in the literature, I expect any bottom-up accountability mechanisms to involve pressuring chiefs using one of those essential elements of chiefs' political status. Consistent with a rational choice framework, I also expect citizens to choose preferably low cost sanctions. The project aims to test the following hypothesis more robustly.

Main hypothesis 1: When undemocratic village leaders do not behave in the interest of the community they represent, citizens sanction them through a variety of social, economic, and political channels, preferably choosing the low cost ones.

The project explores the relationship between councilor pressure and bottom-up pressure by citizens. Similarly, using a rational-choice approach, I expect those pressure to behave as substitute for each other. As councilor pressure was found to be effective in driving and sanctioning chiefs' behavior (Baldwin et al. 2022), when the council of elders sanctions the leader, I expect citizens would refrain from engaging in costly sanctioning action.

Main hypothesis 2: When the council of elders takes an active role in sanctioning the chief, ordinary citizens will be less willing to take an active role in sanctioning the chief. On the contrary, when the council of elders do not take any action, citizens will sanction their chief through a variety of social, economic, and political channels. Horizontal and bottom-up accountability mechanisms would substitute for each other.

The consequences of well-functioning sanctioning institutions

What are the consequences of well-functioning sanctioning institutions on leaders' behavior? When citizens can sanction their chiefs, the relative cost for leaders to undertake actions for their private gains increase. Therefore, it incentivizes pro-social leadership (in the sense of Kosfeld and Rustagi Kosfeld & Rustagi (2015)). Such a mechanism was revealed in a field experiment in Ethiopia by Kahsay and Bulte Kahsay & Bulte (2021) with elected leaders. However, whether a similar mechanism holds with leaders hardly removed from office remains to be seen.

Main hypothesis 3: Villages with a higher ability to sanction their chiefs are associated with a higher responsive leaders.

Main hypothesis 4: Villages with a higher ability to sanction their chiefs are associated with higher leadership quality.

There is currently robust evidence showing the importance of monitoring and sanctioning institutions to maintain ecosystems in good health (Kahsay & Bulte 2021; Ostrom 1990; Walker 2009). However, research primarily focused on elected leaders. The project aims to broaden the scope of the theory to a broader set of bottom-up accountability mechanisms.

Main hypothesis 5: Villages with a higher ability to sanction their chiefs are associated with lower deforestation rates

Secondary hypothesis

I don't expect citizens to have the same ability to sanction their leaders. More specifically, I would expect wealth and gender to be two important characteristics in shaping sanctioning behaviors. The socio-economic conditions influence the perception of how bad a behavior is. For an amount X of money taken from a community project, wealthy citizens will consider it as negligible while the poor will be much more offended. Socialization shapes preferences, values, and norms of correctness. In a society, like rural Sierra Leone, where social spheres are strongly divided between male and female, judgement of behavior correctness defers between male and female. Socialization and differences of judgement will affect the space of legitimize

political actions that a citizen could take. We expect male and the wealthy to have a wider space of legitimize political actions.

Secondary hypothesis 1a: The wealthier the citizen, the more likely he/she use costly sanctioning actions.

Secondary hypothesis 1b: The wealthier the citizen, the more tolerant they are on small corruption behavior.

Secondary hypothesis 2: female citizens are less likely to use costly sanctioning actions

C.3 Empirical strategy

In this section, I develop the empirical strategy and explain the methodology. To test the two first hypotheses, I use two household-level survey experiments. A chief-level survey experiment is also conducted to measure leaders' responsiveness which is essential for testing the third hypothesis. Finally, I analyze advanced satellite images to test the fifth hypothesis.

C.3.1 Data collection and sampling

I aim to test these hypotheses in the Sherbro River Estuary and Yawri Bay in Sierra Leone. The study area is located in the Southern province of Sierra Leone in the district of Bonthe and Moyamba where data is collected in 80 villages. The area is characterized by fishing population with a lack of good arable land for agriculture. Livelihood activities are strongly tied to mangrove forests. Indeed, fishing population smokes fish for conservation that consume a lot of wood. Access to market and cities is very low because of the bad quality of the road network and the high cost of sea transportation.

In each village, 12 households are randomly sampled from the community to participate in the household level survey experiment. The sample is stratified on gender to ensure an equal representation of male and female. Town chief is also selected to participate in the chief level survey experiment.

C.3.2 Three survey experiment designs

Household-level experimental designs

At the household level, I conduct two survey experiments where I vary the nature of the chief behavior. In the first experiment, the scenarios deal with the chief stealing from a community project. In the second experiment, the scenarios deal with the chief allocating land against the community interest. Both survey experiments contains a control scenario and respectively four and three treatment arms. In each village, 12 subjects will be randomly assigned to one of these respectively five or four theoretical scenarios. I use a block randomization on gender to ensure an equal share of male and female in each treatment arms. This strategy is likely to increase the

precision of the estimates and enable to explore heterogeneity effect between male and female (secondary hypothesis 2).

Survey experiment 1

In a village in Sierra Leone, a NGO developed a project for the development of the community. The town chief [*Treatment details*].

- *Pure control: [...] played a key role in managing the project at the village level.*
- *Treatment 1: [...] played a key role in managing the project at the village level. During the project, the town chief took a very small part of money for his own benefits. This treatment arm helps test hypotheses 1.*
- *Treatment 2: [...] played a key role in managing the project at the village level. During the project, the town chief took a very small part of money for his own benefits. The elders in the village went to the chief and sermoned him. This treatment arm is helpful in testing hypothesis 2.*
- *Treatment 3: [...] played a key role in managing the project at the village level. During the project, the town chief took a very small part of money for his own benefits. The elders in the village did not sermoned the chief. This treatment arm is helpful in testing hypothesis 2.*
- *Treatment 4: [...] played a key role in managing the project at the village level. During the project, the town chief took half of the money for his own benefits. This treatment arm is helpful in testing the secondary hypothesis 1.*

After reading the scenario, the subject will answer the following question.

- Question 1: On a 1-5 scale, how much do you agree with the chief's behavior?
- Question 2: Should the citizens of the village take any actions as a reaction? –*i* This measure serves as an indicator about whether the respondents was led
- Question 3: If yes, what actions would you take if you were in a similar situation? (open question asked in a randomly selected set of respondents to reduce survey length)
- Question 4: Some members of the village reacted in this way. In your opinion, is it legitimate? Cross all behavior considered as legitimate in such a situation:
 - *Sanction*₁: Blame the chief directly;
 - *Sanction*₂: Threaten physically the chief;
 - *Sanction*₃: Complain to an elder about the chiefs' behavior;
 - *Sanction*₄: Complain to the paramount chief;
 - *Sanction*₅: Refuse to get married to the chief's member of the family;
 - *Sanction*₇: Refuse to pay local tax;
 - *Sanction*₈: Refuse to participate in farm labor for the chief;
 - *Sanction*₉: Refuse to participate in collective labor;
 - None of the above;
 - Refuse to answer;

- Question 5: In the same list of actions, considering their relative costs, what would you be able to undertake if you were in a similar situation?

Then, we ask two control questions to grasp whether the respondent did understand the scenario.

Survey experiment 2

In a village in Sierra Leone, villagers' livelihood was strongly tied to forest resources. The town chief, in consultation with landowning families, sold [*Treatment details*].

- *Pure control: [...] a small part of the forested land for a conservation project. The land sold could not be used anymore by villagers*
- *Treatment 1: [...] most of the forested land for a conservation project. The land sold could not be used anymore by villagers.* This treatment arm helps test hypotheses 1.
- *Treatment 2: [...] most of the forested land for a conservation project. The land sold could not be used anymore by villagers. The elders in the village went to the chief and sermoned him.* This treatment arm is helpful in testing hypothesis 2.
- *Treatment 3: [. . .] most of the forested land for a conservation project. The land sold could not be used anymore by villagers. The elders in the village did not sermoned the chief.* This treatment arm is helpful in testing hypothesis 2.

After reading the scenario, the subject will answer the following question.

- Question 1: On a 1-5 scale, how much do you agree with the chief's behavior?
- Question 2: Some members of the village reacted in this way. In your opinion, is it legitimate? Cross all behavior considered as legitimate in such a situation:
 - *Sanction₁*: Blame the chief directly;
 - *Sanction₂*: Threaten physically the chief;
 - *Sanction₃*: Complain to an elder about the chiefs' behavior;
 - *Sanction₄*: Complain to the paramount chief;
 - *Sanction₅*: Refuse to get married to the chief's member of the family;
 - *Sanction₇*: Refuse to pay local tax;
 - *Sanction₈*: Refuse to participate in farm labor for the chief;
 - *Sanction₉*: Refuse to participate in collective labor;
 - None of the above;
 - Refuse to answer;

On the survey experiment 2, the number of questions was reduced to minimize the length of the survey. Then, we ask two control questions to grasp whether the respondent did understand the scenario.

C.3.3 The chief level survey experiment

At the village level, I conducted a survey experiment with the chief. The survey experiment is composed of four scenarios. Each town chief faced the four scenarios but with a different order. The pure control scenario always appear first.

- *Pure control: In a village in Sierra Leone, villagers' livelihood was strongly tied to forest resources. A company would like to buy most of the forested land for a conservation project. Villagers could not use anymore the resources.*
- *Treatment 1: Pure control + The community and the paramount chief do support the deal.*
- *Treatment 2: Pure control + The community does support the deal but the paramount chief does not.*
- *Treatment 3: Pure control + The community does not support the deal but the paramount chief does.*

After the scenario was read to the respondent, the enumerator asks the following question: On a 1-5 scale, how likely would you support the deal, if you were in a similar situation?

C.4 Measurement strategy

C.4.1 Measuring the cost of the sanctions listed in the survey experiment

I measure the cost of the sanctions through key informant interviews, and focus group discussions with village members, elders and town chiefs. I categorize the sanctions into two categories: low cost and high cost sanctions.

C.4.2 Individual level measure

The first empirical strategy uses individual-level inference with the following two sanctioning measure and wealth. *Sanctioning measures:* for each participant, I will compute a general sanctioning index (GSI), and a sanctioning index by the relative cost of sanctions (SIC). These two measures are based on the answer to the question: “Some members of the village reacted in this way. In your opinion, is it legitimate? Cross all behavior considered as legitimate in such a situation”.

- General sanctioning index (GSI): for participant j ,

$$GSI_j = \sum_1^9 Sanction_i$$

$Sanction_i$ taking the value 1 if the participant chooses that option and 0 otherwise. The measure are standardized for the purpose of the analysis.

- Sanctioning index by cost (*SIC*): for participant j , and cost c

$$GSI_{j,c} = \sum_{i \in c} Sanction_i$$

$Sanction_i$ of cost c taking the value 1 if the participant chooses that option and 0 otherwise. The measure are standardized for the purpose of the analysis.

Wealth: the wealth of participants will be measured using either a standardized measure of the various assets owned or a standardized measure of the number of farm animals they own that was previously underlined as a relevant measure of wealth.

Controls: those variables will be used to assess the balance of the three experimental groups. Control variables refer to age, gender, number of children, professions, education, wealth, ethnic group, and political and conservation attitudes. In addition, considering the recent development and empirical evidence about the importance of social networks in fostering political participation (Jöst & Lust 2022), I will control for the density of social networks and the priors of the government.

C.4.3 Village level measure

The second empirical strategy uses a village-level inference with the following measures:

Village level accountability index: this is the primary independent variable with theoretical interest. For village k , the bottom-up accountability index (*accindex*) is constructed as the mean of each average *GSI* of each treatment arms.

Leaders' responsiveness: I use the chief-level survey experiment to measure two proxies of the leaders' responsiveness, called *LR*.

- The first proxy measures how much the chief does respond to the community interest when it is in opposition to the paramount chief. To do so, for village j , I measure the difference of support for the deal between treatment 3 and treatment 1:

$$LR_{1j} = Support_{T1} - Support_{T3} \quad (2)$$

- The second proxy measures how much the chief does respond to the community interest compared to the paramount chief interest. To do so, for village j , I measure the difference of support for the deal between treatment 3 and treatment 2:

$$LR_{2j} = Support_{T2} - Support_{T3} \quad (3)$$

$LR1 > 0$ when the chief is responsive to community interest, $LR1 = 0$ when he is not responsive, and $LR1 < 0$ when the chief responds in the contrary to community interest. $LR2 > 0$ when the chief is more responsive to community than the paramount chief, $LR1 = 0$ when the

chief is equally responsive to the community and the paramount chief, and $LR1 < 0$ when the chief is more responsive to the paramount chief than the community.

Leaders' quality: Different measures of leaders' quality is used: a) how much does community members trust the chief, b) the number of annual community meetings, c) the cost for resolving a conflict, d) and how representative leaders chosen by the chief to conduct a land planning activity are. The latter is measured by the number of those leaders selected by the village leader part of its own family.

Village level deforestation index: Deforestation rates from 2015 to 2020 in a 6 km radius from the center will be determined using the Tropical Moist Forest dataset (Vancutsem et al., 2021). Robustness check will include the same measure using buffer zone of 4 to 8km from village center.

Controls: the following control variables are used in the analysis: the size of the village, access to resources, number of ruling families, ties with the paramount chiefs, overall economic condition, inequality between villagers, access to infrastructures, year of leader election

C.5 Analysis

C.5.1 Testing hypothesis 1

I will estimate the effect of leaders' malevolence on attitudes towards the legitimacy of sanctioning behavior with an average treatment effect estimand. If there is covariate balances between the control and the treatment group, I will use the following estimator:

$$GSI_j = \beta_0 + \beta_1 Z_j + \epsilon_j$$

with GSI_j , the outcome variable is the general sanctioning index, β_1 is the Average Treatment Effect, and Z_j is a dummy indicating whether the participant was in the treatment arm 1 or the control group. I will use HC2 standard errors because it is equivalent to a randomization-based Neyman variance estimator (Samii 2023). In addition, I expect the treatment group to have a higher sanctioning index (i.e., finding more sanctions as legitimate behavior). Therefore, I will use a one-tailed test with $H1 > H0$, and $\alpha = 0.05$. As advised in the EGAP modules (source), I will check on the HC2 standard errors by calculating p-values directly using randomization inference, with the difference-in-means as the test statistic.

For each sanction, I will also determine whether the treatment scenario 1 increase the probability to be chosen. I will use a logistic regression framework and determine the odds ratio between the control scenario and the treatment 1 scenario. Using the Wald criteria, I will determine whether the odds ratio is statistically different from 1 at a $\alpha = 0.05$.

$$\text{logit}(Sanction_i) = \beta_0 + \beta_1 Z_j + \epsilon_j$$

C.5.2 Testing hypothesis 2

I will estimate whether horizontal and bottom-up accountability mechanisms work as substitute with an average treatment effect estimand. If there is covariate balance between the treatment arms 1 and 2, I will use the following estimator:

$$GSI_j = \beta_0 + \beta_1 Z_j + \epsilon_j$$

with GSI_j , the outcome variable is the general sanctioning index, β_1 is the Average Treatment effect, and Z_j is a dummy indicating whether the participant was in treatment arm two or treatment arm 1. I will use HC2 standard errors because it is equivalent to a randomization-based Neyman variance estimator (Samii 2023). I expect treatment group 2 to have a lower sanctioning index (i.e., horizontal sanctioning mechanism substitutes bottom-up sanctioning mechanism). Therefore, I will use a one-tailed test with $H1 > H0$, and $\alpha = 0.05$. As advised in the EGAP modules (source), I will check on the HC2 standard errors by calculating p-values directly using randomization inference, with the difference-in-means as the test statistic.

C.5.3 Testing hypothesis 3, 4 and 5

The empirical design does not allow for making any causal claims when testing the hypotheses 3, 4 and 5. Therefore, this section develop a rigorous empirical strategy making statistical association claims. I will use the following OLS regression:

$$Y_k = \beta_0 + \beta_1 accindex_k + \beta_2 X_k \epsilon_k$$

with Y_k being either a proxy of leader responsiveness (for testing hypothesis 3), one of the measure of leaders' quality (for testing hypothesis 4) or the deforestation index measure (for testing hypothesis 5) for village k , β_1 is the effect of the bottom-up accountability index, and X_k are the set of control variables. The estimation strategies consider population size, the number of ruling families, the year of the leader's selection, a measure of inequality, and chiefdom fixed effect as control variables. For hypothesis 3 and 4, I expect villages with better functioning sanctioning institutions (higher bottom-up accountability index) to have more responsive leaders and higher quality leaders. Therefore, I will use a one-tailed test with $H1 > H0$, and $\alpha = 0.05$.

To test hypothesis 5, both a continuous variable and a dummy variable for the deforestation index will be used. I will also add bio-climatic controls to the estimation strategy. I expect villages with better functioning sanctioning institutions to experience lower deforestation rates. Therefore, I will use a one-tailed test with $H1 < H0$, and $\alpha = 0.05$.

C.5.4 Secondary hypotheses

I will estimate whether wealthier citizens and gender tend to sanction more their leaders by using a conditional average treatment effect as stated in (Gerber and Green, 2012). I will use the following estimator:

$$GSI_j = \beta_0 + \beta_1 Z_j + \beta_2 Wealth_j + \beta_3 Z_j Wealth_j + \epsilon_j(9)$$

with GSI_j , the outcome variable is the general sanctioning index for participant j , Z_j a dummy indicating whether the participant was in the treatment arm or the control group, $Wealth_j$ a measure of the wealth of the participant (or gender), and β_3 is the conditional Average Treatment Effect. I will use HC2 standard errors because it is equivalent to a randomization-based Neyman variance estimator (Samii 2023). Furthermore, I expect the wealthier citizens (or male) to have an average treatment effect higher than other citizens. Therefore, I will use a one-tailed test with $H1 > H0$, and $\alpha = 0.05$. As we have two hypotheses in which the estimation uses a covariate-by-treatment interaction, I have a multiple comparisons problem, and I will use the Bonferroni correction using a $\alpha = 0.025$ for both hypotheses. I expect β_1 to be negative for wealthy citizens.

C.5.5 Robustness check

To test the robustness of the analysis for hypotheses 1 to 3, I will first undertake the same analysis with a cluster standard-error at the village level. Then, I will also conduct the estimation separately for the two study areas.

D The study area, measurement strategy and summary statistics

D.1 Study area

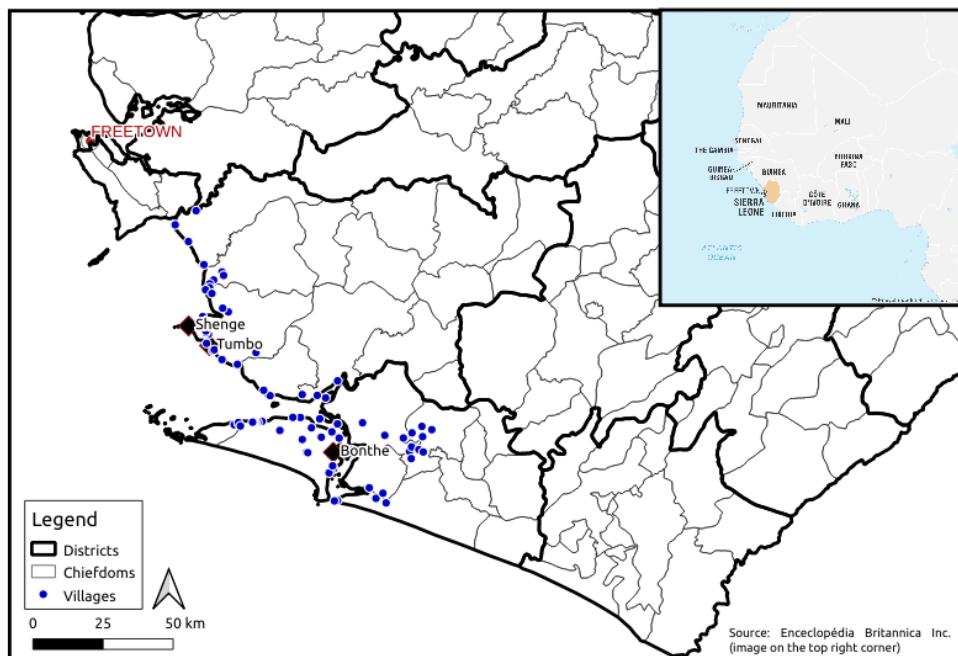


Figure D.1: Map of Sierra Leone, with a focus on the study area, blue dots are the villages sampled.

D.2 Dependent variables used in hypothesis 1 and 2

Table D.1: Table of the measurement strategy for the main dependent variables used to test the hypotheses 1 and 2

Type	Name	Survey question	Range
DV 1	Total number of legitimate sanctions	<i>[After reading one of the five experimental scenario]</i> Some members of the village reacted in this way. In your opinion, is it legitimate? <i>Among the list of nine actions, select all the answers that the respondents considered as legitimate</i>	{0, 1, ..., 8, 9}
DV 2	Number of direct legitimate sanctions	<i>[After reading one of the five experimental scenario]</i> Some members of the village reacted in this way. In your opinion, is it legitimate? <i>Among the list of six actions, select all the answers that the respondents considered as legitimate</i>	{0, 1, ..., 5, 6}
DV 3	Number of indirect legitimate sanctions	<i>[After reading one of the five experimental scenario]</i> Some members of the village reacted in this way. In your opinion, is it legitimate? <i>Among the list of three actions, select all the answers that the respondents considered as legitimate</i>	{0, 1, 2, 3}

D.3 Descriptive statistics of the respondents

Table D.3: Descriptive statistics of the participants and the sampled villages

Variable	N	mean	min	Q1	Q2	Q3	max	NA
Household level variables								
Number of legitimate sanctions	907	3.27	0.00	1.00	4.00	5.00	8.00	0
Number of legitimate direct sanctions	907	1.19	0.00	0.00	1.00	2.00	5.00	0
Number of legitimate indirect sanctions	907	2.08	0.00	1.00	3.00	3.00	3.00	0
Annual income (log)	907	3.64	0.00	3.43	3.72	3.95	4.78	0
Wealth	907	7.66	-4.07	-0.07	5.93	11.93	72.93	0
Age	905	39.64	18.00	30.00	38.00	48.00	85.00	2
Education	890	1.48	0.00	0.00	0.00	0.00	14.00	17
Muslim	907	0.91	0.00	1.00	1.00	1.00	1.00	0
Employed	906	0.13	0.00	0.00	0.00	0.00	1.00	1
Female	907	0.37	0.00	0.00	0.00	1.00	1.00	0
Right to vote	906	0.78	0.00	1.00	1.00	1.00	1.00	1

Notes:

Descriptive statistics (mean, first, second, third quartile, minimum and maximum) of the main variables used to test the hypotheses. The household-level numbers of sanctions are the outcome variables used to test hypotheses 1 and 2. The other household-level variable refers to co-variates.

D.4 Inequality between town chiefs and community members

Table D.4: Citizens and town chiefs' socio-economic characteristics and inequality measures.

	Citizens	Chiefs	Inequality
N	907	76	-
Age	39.64 (12.78)	49.41 (13.94)	12.8 -
Education (in year)	1.48 (3.47)	1.49 (3.39)	1.34 -
Number of children	3.16 (2.2)	4.11 (2.34)	1.28 -
Employed	0.13 (0.33)	0.21 (0.41)	0.18 -
Farm size (in ha)	1.91 (2.32)	2.39 (2.29)	0.76 -
House quality index	1.37 (0.91)	1.25 (1.16)	0.12 -
Material wealth index	1.35 (0.75)	0.97 (0.4)	-0.43 -
Number of livestock	8.52 (8.89)	11.77 (12.95)	5.07 -

Notes:

The citizens and chiefs columns display the average of the socio-economic indicators. Standard errors are in parenthesis. The inequality column represents the average of the difference between the chief indicator and the median of the citizen indicator in the same village.

E Covariate balances

Table E.1: Randomization integrity

	Average					Std. mean diff.		
	C	T1	T2	T3	T4	T1-C	T2-T3	T4-T1
Variable level test								
Wealth	0.09	0.06	-0.03	-0.14	0.01	-0.04	0.11	-0.04
Tenure insecurity	3.19	2.98	3.02	3.24	3.3	-0.21	-0.22	0.32
Income	-2.71	-2.67	-2.69	-2.78	-2.65	0.05	0.09	0.01
Muslim	0.89	0.91	0.9	0.91	0.91	0.02	-0.01	0
Sherbro	0.52	0.44	0.49	0.49	0.51	-0.08	0	0.08
Trust in the chief	3.72	3.56	3.62	3.59	3.63	-0.16**	0.02	0.07
Trust in others	1.31	1.3	1.28	1.29	1.31	-0.01	-0.02	0.01
Employed	0.15	0.08	0.12	0.14	0.13	-0.06*	-0.02	0.05
Female	0.4	0.34	0.38	0.39	0.35	-0.05	0	0.01
Cash emergency	0.69	0.72	0.78	0.74	0.71	0.04	0.04	-0.01
Age	37.75	39.45	43.41	39.46	38.36	1.7	3.95***	-1.09
Education	1.25	1.22	1.63	1.34	1.96	-0.02	0.29	0.74*
Voting rights	0.73	0.74	0.81	0.81	0.8	0.01	0	0.06
Global F test								
F-test						1.61	1.09	1.14
p-value						0.08 ⁺	0.37	0.33

Notes:

This table presents means across treatment arms (columns C to T4) and difference in means and two-sided t-tests for difference in means (columns T1-C to T4-T1) for a set of covariates. P-values are unadjusted for multiple hypotheses testing. A global F-test and p-value are also provided at the bottom of the tables. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ⁺ $p < 0.1$

F Understanding of experimental scenarios

Table F.1: Total number and share of respondents who understood the experimental conditions

	Control	T1	T2	T3	T4	Total
Quantity of money stolen						
	192	170	177	160	162	861
	97 %	96 %	99 %	94 %	88 %	95 %
Behavior of the elders						
	191	161	175	143	168	838
	97 %	91 %	98 %	84 %	91 %	92 %

Notes:

This table displays the number and proportion of respondents who understood the experimental conditions. After answering the outcome questions, respondents were asked to recall the amount of money taken by the town chief, if any, and the behavior of the elders, if any.

G Deviation from the pre-analysis plan

This section outlines the deviations from the pre-analysis plan and provides justifications for each based on the reporting table chart. The deviations are categorized into three subsections: (a) sampling and experimental conditions, (b) hypothesis wording, (c) measurement strategy, and (d) hypothesis testing. All modifications were made solely to enhance the precision and clarity of the analysis. Additionally, all pre-registered tests have been conducted and are reported at the end of this section, aligning with the arguments and findings presented in the main text. None of them contradicts the findings reported in the main text.

G.1 Sampling and experimental conditions

In the registered pre-analysis plan, the experiment included two scenarios: a community project and a land deal scenario. For the latter, a crucial observation made during the fieldwork stage was the inconsistency between the scenario presented and the prevailing contextual realities. Specifically, it was determined that the sale of lands, which formed the basis of the scenario, does not align with customary practices (lands can only be leased) and land transactions typically involve consultation with the paramount chief. Second, the scenario lacks a pure control, thus limiting the ability to isolate and assess the specific effects of the variables under investigation. As a result, in adherence to the registered pre-analysis plan, the detailed results of the land deal scenario are presented in Appendix G.6. The results align with our expectations and are consistent with those presented in the main text.

Table G.1: Deviation from the pre-analysis plan regarding sampling strategy and experimental conditions

	Specified in pre-registration? (Yes, no)	Reported in manuscript? (Yes, page(s) in main text; Yes, page(s) in appendix; No)	Deviations (justification and loca- tion in the manuscript)
Sampling	Yes	Yes, p. 9 in main text	None
Sample Exclusions	No	No	Individuals who were unavailable for the survey were replaced using a randomly sampled list of pre- identified replacements.
Experimental Conditions	5 experimental conditions for survey experiment 1; 4 condi- tions for survey experiment 2	Only survey experiment 1 re- ported in main text (p. 11); sur- vey experiment 2 in Appendix (p. 56)	The pre-analysis plan included two scenarios: a community project and a land deal. During fieldwork, the land deal scenario was found to misalign with contextual realities—lands can only be leased, not sold, and transactions often involve the paramount chief. Additionally, the scenario lacked a pure control, limiting causal inference. In line with the pre-analysis plan, results for the land deal are presented in Appendix G.6 and are consistent with the main findings.

G.2 Inclusion and wording of the hypotheses and deviations

Table G.2: Deviation from the pre-analysis plan regarding the wording of the hypotheses and their justifications.

	Pre-registered hypotheses	New wording	Justifications
Hyp. 1	<i>When undemocratic village leaders do not behave in the interest of the community they represent, citizens sanction them through a variety of social, economic, and political channels, preferably choosing the low cost ones</i>	Hyp. 1: Village members view sanctions against chiefs who violate community interests as legitimate and feasible. (page 7 in the main text)	Pre-registered Hypothesis 1 includes two separate assertions, each of which requires independent testing. The second assertion can not be tested experimentally and is dropped from the analysis with descriptive statistics provided in Appendix K. Additionally, we simplified the wording to enhance clarity and comprehension. The content and directional predictions of the hypothesis remain unchanged.
Hyp. 2	<i>When the council of elders takes an active role in sanctioning the chief, ordinary citizens will be less willing to take an active role in sanctioning the chief. On the contrary, when the council of elders do not take any action, citizens will sanction their chief through a variety of social, economic, and political channels. Horizontal and bottom-up accountability mechanisms would substitute for each other.</i>	Hyp. 2: community members will be less willing to view sanction as legitimate and feasible if the council of elders already blamed the chief for their action. (page 8 in the main text)	Pre-registered Hypothesis 2 includes unnecessary wording and confusing languages. We simplified the wording to enhance clarity and comprehension. The content and directional predictions of the hypothesis remain unchanged.
Hyp. 3	<i>Villages with a higher ability to sanction their chiefs are associated with higher responsive leaders.</i>	Not included in the paper	Pre-registered Hypothesis 3 uses non experimental evidence and is beyond the scope of the paper.
Hyp. 4	<i>Villages with a higher ability to sanction their chiefs are associated with higher leadership quality</i>	Not included in the paper	Pre-registered Hypothesis 4 uses non experimental evidence and is beyond the scope of the paper.
Hyp. 5	<i>Villages with a higher ability to sanction their chiefs are associated with lower deforestation rates</i>	Not included in the paper	Pre-registered Hypothesis 5 uses non experimental evidence and is beyond the scope of the paper.

G.3 Measurement analysis and deviations

Table G.3: Deviation from the pre-analysis plan regarding the measurement strategy and their justifications

Pre-registered measurement		Deviations	Justifications
INDEPENDENT VARIABLES			
Hyp. 1	Treatment 1 vs. Control	<i>None</i>	
Hyp. 2	Treatment 2 vs. Treatment 3	<i>None</i>	
DEPENDENT VARIABLES			
Hyp. 1	General sanctioning index and a Sanctioning index by costs (high vs. low)	General sanctioning index (p9), sanctioning index by type (indirect vs. direct, results reported in p12) and sanctioning index by costs (high vs. low, results reported in Appendix K)	We introduce the distinction between indirect and direct sanctions, as they are theoretically significant and should be conceptually separated.
Hyp. 2	General sanctioning index and a Sanctioning index by costs (high vs. low)	General sanctioning index, sanctioning index by type (indirect vs. direct).	We introduce the distinction between indirect and direct sanctions, as they are theoretically significant and should be conceptually separated.

G.4 Empirical strategy and deviations

Table G.4: Deviation from the pre-analysis plan regarding the measurement strategy and their justifications

	Pre-registered specification	Deviations	Justifications
Hyp. 1	$GSI_j = \beta_0 + \beta_1 Z_j + \epsilon_j$	$GSI_j = \beta_0 + \beta_1 Z_j + \gamma_g + \gamma_v + \epsilon_j$	We add experimental block fixed effects, procedure called precision weighting. Gerber & Green (2012) show that such a procedure increases precision. As our sample size is somewhat limited, below 1000 respondents for 5 experimental conditions, it is likely to be a more efficient strategy. The pre-registered strategy is also reported in the Appendix and yield to similar result.
Hyp. 2	$GSI_j = \beta_0 + \beta_1 Z_j + \epsilon_j$	$GSI_j = \beta_0 + \beta_1 Z_j + \gamma_g + \gamma_v + \epsilon_j$	We add experimental block fixed effects, procedure called precision weighting. Gerber & Green (2012) show that such a procedure increases precision. As our sample size is somewhat limited, below 1000 respondents for 5 experimental conditions, it is likely to be a more efficient strategy. The pre-registered strategy is also reported in the Appendix and yield to similar result.

G.5 Pre-Registered and Main Text Results Overview

Table G.5: Table of the main deviation from the pre-analysis plan. The pre-registered hypotheses are laid out with their results using the specification described in the pre-analysis plan.

Pre-registered hypotheses	Test	PAP	p-values Main text specification
H1: When undemocratic village leaders do not behave in the interest of the community they represent, citizens sanction them through a variety of social, economic, and political channels, preferably choosing the low cost ones	$H_1 > H_0$	$p < .01$	$p < .01$
H2: When the council of elders takes an active role in sanctioning the chief, ordinary citizens will be less willing to take an active role in sanctioning the chief. On the contrary, when the council of elders do not take any action, citizens will sanction their chief through a variety of social, economic, and political channels. Horizontal and bottom-up accountability mechanisms would substitute for each other.	$H_1 < H_0$	$p < .95$	$p < .99$

G.6 Main results using the land deal survey experiment

In the registered pre-analysis plan, two scenarios were mentioned: a community project and a land deal scenario. For the latter, a crucial observation made during the fieldwork stage was the inconsistency between the scenario presented and the prevailing contextual realities. Specifically, it was determined that the sale of lands, which formed the basis of the scenario, does not align with customary practices (lands can only be leased) and land transactions typically involve consultation with the paramount chief. Second, the scenario lacks a pure control, thus limiting the ability to isolate and assess the specific effects of the variables under investigation. As a result, in adherence to the registered pre-analysis plan, the detailed results of the land deal scenario are presented in the appendices for reference. The results remain consistent with the one reported in the main text. More specifically, the scenarios were as followed:

- **Control arm:** *In a village in Sierra Leone, villagers' livelihood was strongly tied to forest resources. The town chief, in consultation with landowning families, sells a small part of the forested land for a conservation project. The land sold could not be used anymore by villagers.*
- **Treatment 1:** *In a village in Sierra Leone, villagers' livelihood was strongly tied to forest resources. The town chief, in consultation with landowning families, sells most of the forested land for a conservation project. The land sold could not be used anymore by villagers.*
- **Treatment 2:** *In a village in Sierra Leone, villagers' livelihood was strongly tied to forest resources. The town chief, in consultation with landowning families, sells most of the forested land for a conservation project. The land sold could not be used anymore by villagers. The elders in the village went to the chiefs and sermoned him.*
- **Treatment 3:** *In a village in Sierra Leone, villagers' livelihood was strongly tied to forest resources. The town chief, in consultation with landowning families, sells most of the forested land for a conservation project. The land sold could not be used anymore by villagers. The elders in the village did not sermon the chief.*

Test for hypothesis 1

Table G.6: Average treatment effects of chief malevolance on the total number of legitimate sanctions, of direct and indirect sanctions during the land deal survey experiment

	Number of legitimate sanction					
	General		Direct		Indirect	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.58*** (0.17)	0.63*** (0.17)	0.41*** (0.10)	0.45*** (0.09)	0.17 (0.11)	0.19+ (0.10)
Control mean	3.18	3.18	0.991	0.991	2.19	2.19
DV range	{0, 9}		{0, 6}		{0, 3}	
Block FE	No	Yes	No	Yes	No	Yes
Adj. R ²	0.02	0.14	0.04	0.18	0.00	0.13
Num. obs.	441	441	441	441	441	441

Notes:

Three outcome variable are used: the total sum of sanctions considered as legitimate (column 1 and 2), the total sum of direct sanctions considered as legitimate (column 3 and 4) and the total sum of indirect sanctions considered as legitimate (column 5 and 6). Treatment refers to a dummy comparing the scenario where the town chief sells half of the forested land for a conservation project compared to a scenario where the chief sells only a small part. The specification from the pre-analysis plan is used in column 1, 3, and 5. Village and block fixed effects are added in column 2, 4, and 6. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$

Test of the hypothesis 2

Table G.7: Average treatment effects of elders blaming the town chief on the total number of legitimate sanctions, of direct and indirect sanctions during the land deal survey experiment

	Number of legitimate sanction					
	General		Direct		Indirect	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.15 (0.16)	0.15 (0.16)	0.01 (0.09)	0.01 (0.09)	0.13 (0.10)	0.14 (0.10)
Control mean	3.69	3.69	1.28	1.28	2.42	2.42
DV range	{0, 9}		{0, 6}		{0, 3}	
Block FE	No	Yes	No	Yes	No	Yes
Adj. R ²	-0.00	0.00	-0.00	0.05	0.00	0.04
Num. obs.	413	413	413	413	413	413

Notes:

Three outcome variable are used: the total sum of sanctions considered as legitimate (column 1 and 2), the total sum of direct sanctions considered as legitimate (column 3 and 4) and the total sum of indirect sanctions considered as legitimate (column 5 and 6). Treatment refers to a dummy comparing the scenario where the town chief sells half of the forested land for a conservation project and the elders do blame the chief compared to a scenario where the chief sells half of the forested land and the elders do not blame the chief. The specification from the pre-analysis plan is used in column 1, 3, and 5. Village and block fixed effects are added in column 2, 4, and 6. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$

H Open-question outcome summary

Table H.1: Categorized Responses to the Open Question - What Would You Do in a Similar Situation?

Variable	N	Frequency	Percentage
Collective Actions			
Mobilize citizens	404	13	3 %
Town meetings	404	5	1 %
Direct actions			
Force give back money	404	80	20 %
Blame directly	404	49	12 %
Refuse to take orders	404	17	4 %
Vote against him	404	5	1 %
Attack the chief	404	4	1 %
Remove him from the project	404	3	1 %
Stop working for him	404	2	0 %
Fine him	404	1	0 %
Indirect actions			
Report to elders	404	97	24 %
Report to paramount chief	404	55	14 %
Report to section chief	404	53	13 %
Report to higher authorities	404	48	12 %
Report to village authorities	404	41	10 %
Ask for suspension	404	40	10 %
Report to police	404	16	4 %
Bring case to court	404	10	2 %
Report to NGO	404	8	2 %
No actions			
No actions	404	71	18 %
Preference for the chief giving back the money	404	37	9 %
Not understandable	404	1	0 %

Notes:

For half of the sample (N=404), after receiving the scenario, we asked an open question: What would you do in a similar situation? We categorized the actions into the categories reported in the table with the absolute and relative frequency.

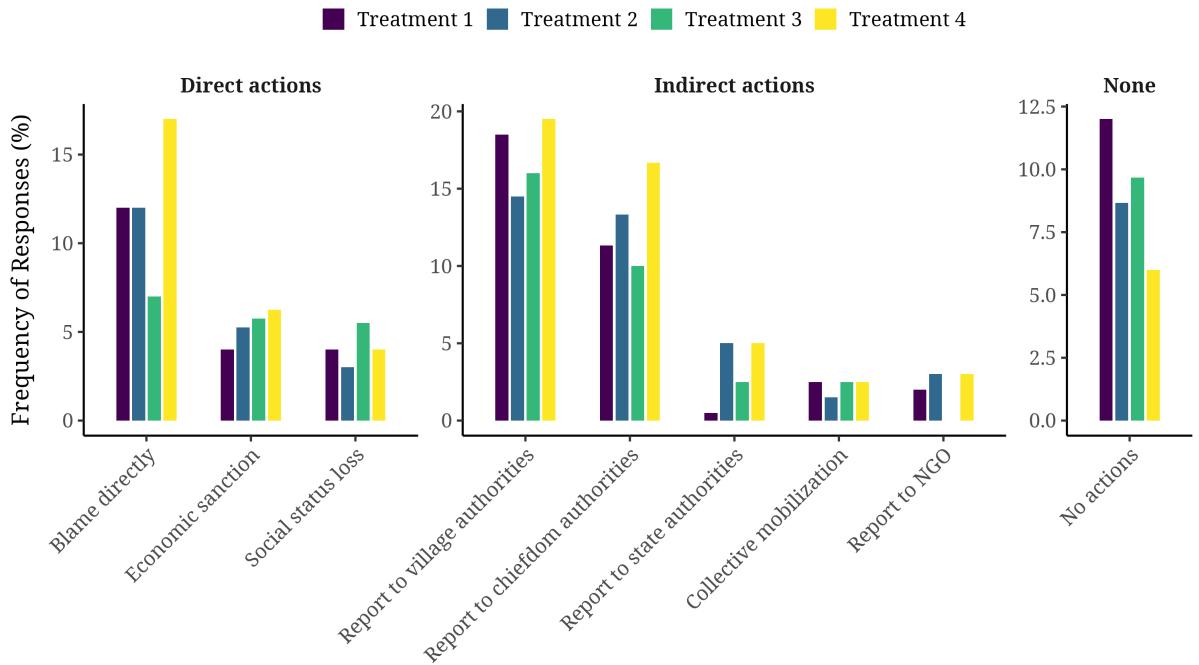


Figure H.1: Categorized Responses grouped to the Open Question ‘What Would You Do in a Similar Situation?’ by scenario

I Treatment effects on the main outcomes

I.1 Outcome 1: disagreement with the chief's behavior

Effect of the gravity of the theft

Table I.1: Average treatment effects of chief malevolence on disagreement with the chief's behavior

Treatment	Disagreement with the chief's behavior					
	T1 - C			T4 - T1		
	0.7*** (0.03)	0.7*** (0.03)	0.7*** (0.03)	0.1*** (0.02)	0.09*** (0.02)	0.09*** (0.02)
p-value	0.028	< 0.001	< 0.001	0.022	< 0.001	< 0.001
p-bonferroni	0.169	< 0.001	< 0.001	0.131	< 0.001	< 0.001
p-fdr	0.028	< 0.001	< 0.001	0.026	< 0.001	< 0.001
Control mean	0.13	0.13	0.13	0.83	0.83	0.83
DV range	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]
Block FE	No	Yes	Yes	No	Yes	Yes
Observations	374	374	374	361	361	361
R ²	0.62	0.68	0.68	0.05	0.27	0.27

Note:

The outcome variable refers to a 1-5 scale about how much citizens disagree with chiefs' behavior (standardized into a 0 to 1 variable). The table shows the analysis comparing the treatment 1 with the control group in columns 1 to 3, and comparing treatment 4 to the treatment 1 in column 4 to 6. The treatment 1 refers to a scenario where the chief takes a small part of the money for his own benefit. The treatment 4 refers to a scenario where the chief takes half of the money for his own benefit. Robust standard errors in parenthesis. Column 1 and 4 refers to the empirical specification from the pre-analysis plan. Block fixed effects are used in column 2, 3, 5 and 6. Robust cluster standard errors at the village level are used in column 3 and 6. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

Effect of the council of elders' behavior

Table I.2: Average treatment effects of council of elders' behavior on disagreement with the chief's behavior

	Disagreement with the chief's behavior					
	T2 - T3			T2 - T1		
	Treatment	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0.04 ⁺ (0.02)	0.04 ⁺ (0.02)
p-value	0.023	0.377	0.336	0.024	0.064	0.056
p-bonferroni	0.141	1	1	0.145	0.383	0.336
p-fdr	0.072	0.377	0.377	0.072	0.096	0.096
Control mean	4.43	4.43	4.43	4.32	4.32	4.32
DV range	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]
Block FE	No	Yes	Yes	No	Yes	Yes
Observations	349	349	349	355	355	355
R ²	< 0.001	0.22	0.22	0.01	0.21	0.21

Note:

The outcome variable refers to a 1-5 scale about how much citizens disagree with chiefs' behavior (standardized into a 0 to 1 variable). The table shows the analysis comparing the treatment 2 with the treatment 3 in columns 1 to 3, and comparing treatment 2 to the treatment 1 in column 4 to 6. The treatment 2 refers to a scenario where the chief takes a small part of the money for his own benefit and the elders blamed the chief, while in treatment 3 they did not. The treatment 1 refers to a scenario where the chief takes half of the money for his own benefit and no mention of the elders. Robust standard errors in parenthesis. Column 1 and 4 refers to the empirical specification from the pre-analysis plan. Block fixed effects are used in column 2, 3, 5 and 6. Robust cluster standard errors at the village level are used in column 3 and 6. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

I.2 Outcome 2: Citizens should take actions against the chief

Effect of the gravity of the theft

Table I.3: Average treatment effects of chief malevolence on whether citizens should take any actions

	Citizens should take any action					
	T1 - C			T4 - T1		
Treatment	0.73*** (0.04)	0.74*** (0.03)	0.74*** (0.04)	0.12*** (0.04)	0.12*** (0.03)	0.12*** (0.03)
p-value	0.035	< 0.001	< 0.001	0.036	< 0.001	0.001
p-bonferroni	0.213	< 0.001	< 0.001	0.219	0.001	0.003
p-fdr	0.036	< 0.001	< 0.001	0.036	< 0.001	0.001
Control mean	0.07	0.07	0.07	0.8	0.8	0.8
DV range	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
Block FE	No	Yes	Yes	No	Yes	Yes
Observations	374	374	374	361	361	361
R ²	0.54	0.62	0.62	0.03	0.27	0.27

Note:

The outcome variable refers to a dummy variable indicating whether the respondents consider citizens should take any action given the hypothetical scenario. The table shows the analysis comparing the treatment 1 with the control group in columns 1 to 3, and comparing treatment 4 to the treatment 1 in column 4 to 6. The treatment 1 refers to a scenario where the chief takes a small part of the money for his own benefit. The treatment 4 refers to a scenario where the chief takes half of the money for his own benefit. Robust standard errors in parenthesis. Column 1 and 4 refers to the empirical specification from the pre-analysis plan. Block fixed effects are used in column 2, 3, 5 and 6. Robust cluster standard errors at the village level are used in column 3 and 6. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

Effect of the council of elders' behavior

Table I.4: Average treatment effects of the council of elders' behavior on whether citizens should take any actions

	Citizens should take any action					
	T2 - T3			T2 - T1		
	Treatment	0.03 (0.04)	0.03 (0.03)	0.03 (0.03)	0.09* (0.04)	0.08* (0.03)
p-value	0.036	0.277	0.184	0.039	0.021	0.017
p-bonferroni	0.214	1	1	0.231	0.128	0.102
p-fdr	0.058	0.277	0.221	0.058	0.058	0.058
Control mean	0.86	0.86	0.86	0.8	0.8	0.8
DV range	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
Block FE	No	Yes	Yes	No	Yes	Yes
Observations	349	349	349	355	355	355
R ²	< 0.001	0.27	0.27	0.01	0.25	0.25

Note:

The outcome variable refers to a dummy variable indicating whether the respondents consider citizens should take any action given the hypothetical scenario. The table shows the analysis comparing the treatment 2 with the treatment 3 in columns 1 to 3, and comparing treatment 2 to the treatment 1 in column 4 to 6. The treatment 2 refers to a scenario where the chief takes a small part of the money for his own benefit and the elders blamed the chief, while in treatment 3 they did not. The treatment 1 refers to a scenario where the chief takes half of the money for his own benefit and no mention of the elders. Robust standard errors in parenthesis. Column 1 and 4 refers to the empirical specification from the pre-analysis plan. Block fixed effects are used in column 2, 3, 5 and 6. Robust cluster standard errors at the village level are used in column 3 and 6. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

I.3 Legitimacy of the sanctions

Effect of the gravity of the theft

Table I.5: Average treatment effects of chief malevolence on the total number of legitimate sanctions, the number of direct and indirect legitimate sanctions against the village chief

	Number of legitimate sanctions (sum)					
	T1 - C			T4 - T1		
	General	Direct	Indirect	General	Direct	Indirect
Treatment	3.04*** (0.16)	1.15*** (0.09)	1.89*** (0.1)	0.54*** (0.14)	0.27** (0.1)	0.27*** (0.08)
p-value	< 0.001	< 0.001	< 0.001	< 0.001	0.007	0.001
p-bonferroni	< 0.001	< 0.001	< 0.001	0.001	0.044	0.005
p-fdr	< 0.001	< 0.001	< 0.001	< 0.001	0.007	0.001
Control mean	0.65	0.13	0.52	3.64	1.25	2.4
DV range	{0, 9}	{0, 6}	{0, 3}	{0, 9}	{0, 6}	{0, 3}
Observations	374	374	374	361	361	361
R ²	0.54	0.38	0.57	0.44	0.34	0.29

Note:

Three outcome variables are used: general, direct and indirect. General is the total sum of sanctions considered as legitimate. Direct is the sum of sanctions considered as legitimate targeting directly the chief. Indirect is the sum of sanctions considered as legitimate targeting a higher authority. T1-C tests Treatment 1 where the quantity of money stolen is very small vs the control group, and T4-T1 tests Treatment 4, where half of the money is stolen vs Treatment 1 (hypothesis 1). Village and block fixed effects are used. Robust standard errors in parenthesis. Reported p-values include the unadjusted p-value (p-value), Bonferroni-adjusted p-value (p-bonferroni), and False Discovery Rate-adjusted p-value using the Benjamini-Hochberg method (p-fdr). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

Effect of the council of elders' behavior

Table I.6: Average treatment effects of elders' behavior on the total number of legitimate sanctions, the number of direct and indirect legitimate sanctions against the village chief

	Number of legitimate sanctions (sum)					
	T2 - T3			T2 - T1		
	General	Direct	Indirect	General	Direct	Indirect
Treatment	0.37* (0.15)	0.13 (0.12)	0.24** (0.09)	0.38** (0.15)	0.24* (0.1)	0.14 (0.09)
p-value	0.017	0.258	0.007	0.009	0.019	0.108
p-bonferroni	0.103	1	0.042	0.057	0.111	0.65
p-fdr	0.028	0.258	0.028	0.028	0.028	0.13
Control mean	3.81	1.44	2.36	3.64	1.25	2.4
DV range	{0, 9}	{0, 6}	{0, 3}	{0, 9}	{0, 6}	{0, 3}
Observations	349	349	349	355	355	355
R ²	0.34	0.18	0.29	0.42	0.36	0.31

Note:

Three outcome variables are used: general, direct and indirect. General is the total sum of sanctions considered as legitimate. Direct is the sum of sanctions considered as legitimate targeting directly the chief. Indirect is the sum of sanctions considered as legitimate targeting an higher authority. T2-T3 tests treatment 2 (where the elders do blame the chief) vs treatment 3 (where the elders do not blame the chief), and T1-T3 tests treatment 1 (which made no mention of elder behavior) vs treatment 3 (where the elders do not blame the chief). Control mean refers to the average number of sanctions considered as legitimate for the scenario of reference (treatment 3 and treatment 1). Village and block fixed effects are used. Robust standard errors in parenthesis. Reported p-values include the unadjusted p-value (p-value), Bonferroni-adjusted p-value (p-bonferroni), and False Discovery Rate-adjusted p-value using the Benjamini-Hochberg method (p-fdr). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

Outcome by sanction and treatment condition

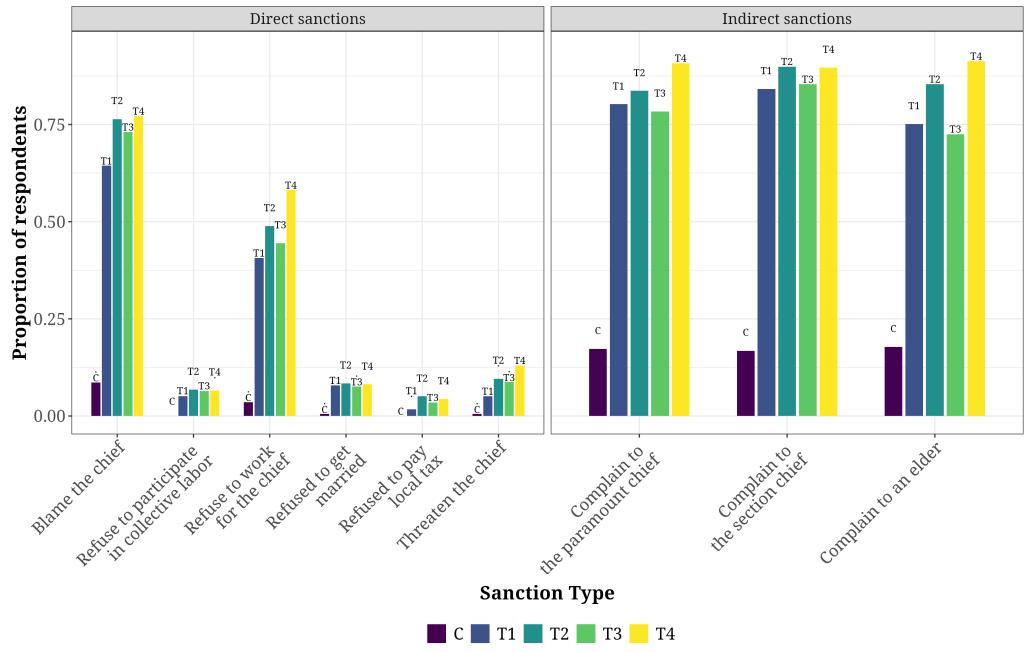


Figure I.1: Proportion of respondents considering each sanction as legitimate by treatment condition

I.4 Feasibility of the sanctions

Effect of the gravity of the theft

Table I.7: Average treatment effects of chief malevolence on the total number of feasible sanctions, the number of direct and indirect feasible sanctions against the village chief

	Number of feasible sanctions (sum)					
	T1 - C			T4 - T1		
	General	Direct	Indirect	General	Direct	Indirect
Treatment	2.92*** (0.17)	1.21*** (0.09)	1.69*** (0.11)	0.37* (0.17)	0.24* (0.11)	0.15 (0.1)
p-value	< 0.001	< 0.001	< 0.001	0.028	0.034	0.145
p-bonferroni	< 0.001	< 0.001	< 0.001	0.169	0.201	0.868
p-fdr	< 0.001	< 0.001	< 0.001	0.04	0.04	0.145
Control mean	0.63	0.12	0.51	3.52	1.31	2.19
DV range	{0, 9}	{0, 6}	{0, 3}	{0, 9}	{0, 6}	{0, 3}
Observations	373	374	374	360	361	361
R ²	0.49	0.37	0.52	0.32	0.29	0.21

Note:

Three outcome variables are used: general, direct and indirect. General is the total sum of sanctions considered as feasible. Direct is the sum of sanctions considered as feasible targeting directly the chief. Indirect is the sum of sanctions considered as feasible targeting a higher authority. T1-C tests Treatment 1 where the quantity of money stolen is very small vs the control group, and T4-T1 tests Treatment 4, where half of the money is stolen vs Treatment 1 (hypothesis 1). Village and block fixed effects are used. Robust standard errors in parenthesis. Reported p-values include the unadjusted p-value (p-value), Bonferroni-adjusted p-value (p-bonferroni), and False Discovery Rate-adjusted p-value using the Benjamini-Hochberg method (p-fdr). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

Effect of the council of elders' behavior

Table I.8: Average treatment effects of elders' behavior on the total number of feasible sanctions, the number of direct and indirect feasible sanctions against the village chief

	Number of feasible sanctions (sum)					
	T2 - T3			T2 - T1		
	General	Direct	Indirect	General	Direct	Indirect
Treatment	0.42* (0.18)	0.06 (0.12)	0.36*** (0.1)	0.5** (0.18)	0.25* (0.12)	0.27** (0.1)
p-value	0.021	0.615	< 0.001	0.006	0.038	0.006
p-bonferroni	0.124	1	0.002	0.038	0.229	0.038
p-fdr	0.031	0.615	0.002	0.013	0.046	0.013
Control mean	3.65	1.5	2.16	3.52	1.31	2.19
DV range	{0, 9}	{0, 6}	{0, 3}	{0, 9}	{0, 6}	{0, 3}
Observations	349	349	349	354	355	355
R ²	0.24	0.18	0.24	0.28	0.23	0.26

Note:

Three outcome variables are used: general, direct and indirect. General is the total sum of sanctions considered as feasible. Direct is the sum of sanctions considered as feasible targeting directly the chief. Indirect is the sum of sanctions considered as feasible targeting an higher authority. T2-T3 tests treatment 2 (where the elders do sermon the chief) vs treatment 3 (where the elders do not sermon the chief), and T1-T3 tests treatment 1 (which made no mention of elder behavior) vs treatment 3 (where the elders do not sermon the chief). Control mean refers to the average number of sanctions considered as feasible for the scenario of reference (treatment 3 and treatment 1). Village and block fixed effects are used. Robust standard errors in parenthesis. Reported p-values include the unadjusted p-value (p-value), Bonferroni-adjusted p-value (p-bonferroni), and False Discovery Rate-adjusted p-value using the Benjamini-Hochberg method (p-fdr). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

J Robustness check for the statistical test of hypotheses 1 and 2

J.1 Comparison of treatment 1 with the control arm

Table J.1: Robustness check for hypothesis 1 comparing the treatment 1 with the control arm. Treatment effects on the total number of legitimate sanctions, the number of direct and indirect legitimate sanctions.

<i>Outcome: total sum of legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	2.99*** (0.17)	3.04*** (0.23)	3.07*** (0.23)	3.14*** (0.24)
Control mean	0.65	0.65	0.65	0.59
Adj. R ²	0.46	0.54	0.54	0.57
Number of observations	374	374	374	362
<i>Outcome: sum of direct legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	1.12*** (0.09)	1.15*** (0.12)	1.12*** (0.12)	1.20*** (0.12)
Control mean	0.13	0.13	0.13	0.10
Adj. R ²	0.31	0.38	0.39	0.41
Number of observations	374	374	374	362
<i>Outcome: sum of indirect legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	1.88*** (0.11)	1.89*** (0.14)	1.95*** (0.13)	1.93*** (0.14)
Control mean	0.52	0.52	0.52	0.49
Adj. R ²	0.43	0.57	0.59	0.59
Number of observations	374	374	374	362
Block FE	No	Yes	Yes	Yes
Cluster SE	No	Yes	Yes	Yes
Covariates	No	No	Yes	No
Observation dropped	No	No	No	Yes

Notes:

Treatment tests treatment 1 against the control group with the treatment 1 being the scenario where the chiefs take a small part of the money for its own benefit. Four robustness check strategies are used. The first column is the PAP empirical strategy without block fixed effects. The second column uses cluster robust standard error at the village level. The third column adds covariates that were not balanced (employment rates and trust to the chief), and the fourth column excludes respondents that did not understand the scenario. Village and block fixed effects are used in the last three columns. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

J.2 Comparison of treatment 4 with the treatment 1

Table J.2: Robustness check for hypothesis 1 comparing the treatment 4 with the treatment 1. Treatment effects on the total number of legitimate sanctions, the number of direct and indirect legitimate sanctions.

<i>Outcome: total sum of legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	0.75*** (0.18)	0.54*** (0.13)	0.49*** (0.13)	0.70*** (0.14)
Control mean	3.64	3.64	3.64	3.66
Adj. R ²	0.05	0.44	0.43	0.45
Number of observations	361	361	353	332
<i>Outcome: sum of direct legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	0.43*** (0.12)	0.27* (0.10)	0.24* (0.10)	0.33** (0.11)
Control mean	1.25	1.25	1.25	1.26
Adj. R ²	0.03	0.34	0.33	0.35
Number of observations	361	361	353	332
<i>Outcome: sum of indirect legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	0.32*** (0.09)	0.27*** (0.08)	0.25** (0.08)	0.37*** (0.08)
Control mean	2.40	2.40	2.40	2.39
Adj. R ²	0.03	0.29	0.30	0.33
Number of observations	361	361	353	332
Block FE	No	Yes	Yes	Yes
Cluster SE	No	Yes	Yes	Yes
Covariates	No	No	Yes	No
Observation dropped	No	No	No	Yes

Notes:

Treatment tests treatment 4 against the treatment 1 with the treatment 4 being the scenario where the chiefs take half of the money for its own benefit and treatment 1 where the chiefs take only a small part of the money of its own benefit. Four robustness check strategies are used. The first column is the PAP empirical strategy without block fixed effects. The second column uses cluster robust standard error at the village level. The third column adds covariates that were not balanced (education level), and the fourth column excludes respondents that did not understand the scenario. Village and block fixed effects are used in the last three columns. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

J.3 Sensitivity analysis to experimenter demand effect (EDE) - hypothesis 1

de Quidt et al. (2018) bound the experimenter demand effect from 0.1 to 0.3 standard deviation of the outcome variable. As a consequence, we used three sensitivity analysis. We standardized the outcome variable and reduce of respectively 0.1, 0.2, and 0.3 the measured outcome in the treatment group to account for potential experimenter demand effect.

Table J.3: Sensitivity analysis to different strength of experimenter demand effect (EDE) for hypothesis 1 comparing the treatment 1 with the control arm.

	<i>Outcome: total sum of legitimate sanctions</i>			
	(1)	(2)	(3)	(4)
Treatment	1.41*** (0.11)	1.31*** (0.11)	1.21*** (0.11)	1.11*** (0.11)
Strength of the EDE (in std. dev.)	0	0.1	0.2	0.3
Adj. R ²	0.54	0.51	0.48	0.45
Num. obs.	374	374	374	374

Notes:

Treatment tests treatment 1 against the control group with the treatment 1 being the scenario where the chiefs take a small part of the money for its own benefit. Three sensitivity analysis are used (column 2 to 4). The first column is the main empirical strategy with the outcome variable being standardized. The second, third and fourth column use an experimenter demand effect of 0.1, 0.2, and 0.3 standard deviation respectively. Cluster robust standard error at the village level. Village and block fixed effects are used in the last three columns.
 *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

J.4 Robustness check for the statistical test of hypotheses 2

Table J.4: Robustness check for hypothesis 2 comparing the treatment 2 with the treatment 3. Treatment effects on the total number of legitimate sanctions, the number of direct and indirect legitimate sanctions.

<i>Outcome: total sum of legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	0.33 ⁺ (0.18)	0.37* (0.17)	0.35* (0.18)	0.48** (0.17)
Control mean	3.81	3.81	3.81	3.84
Adj. R ²	0.01	0.34	0.34	0.37
Number of observations	349	349	348	310
<i>Outcome: sum of direct legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	0.11 (0.12)	0.13 (0.13)	0.12 (0.13)	0.19 (0.14)
Control mean	1.44	1.44	1.44	1.44
Adj. R ²	0.00	0.18	0.18	0.18
Number of observations	349	349	348	310
<i>Outcome: sum of indirect legitimate sanctions</i>				
	(1)	(2)	(3)	(4)
Treatment	0.23* (0.10)	0.24** (0.09)	0.24* (0.09)	0.29** (0.09)
Control mean	2.36	2.36	2.36	2.40
Adj. R ²	0.01	0.29	0.29	0.35
Number of observations	349	349	348	310
Block FE	No	Yes	Yes	Yes
Cluster SE	No	Yes	Yes	Yes
Covariates	No	No	Yes	No
Observation dropped	No	No	No	Yes

Notes:

Treatment tests treatment 2 against the treatment 3. The treatment 2 is a scenario where the elders did blame the town chief. The treatment 3 is a scenario where the elders did not blame the town chief. Four robustness check strategies are used. The first column is the PAP empirical strategy without block fixed effects. The second column uses cluster robust standard error at the village level. The third column adds covariates that were not balanced (education level), and the fourth column excludes respondents that did not understand the scenario. Village and block fixed effects are used in the last three columns. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

J.5 Sensitivity analysis to experimenter demand effect (EDE) - hypothesis 2

de Quidt et al. (2018) bound the experimenter demand effect from 0.1 to 0.3 standard deviation of the outcome variable. As a consequence, we used three sensitivity analysis. We standardized the outcome variable and reduce of respectively 0.1, 0.2, and 0.3 the measured outcome in the treatment group to account for potential experimenter demand effect.

Table J.5: Sensitivity analysis to different strength of experimenter demand effect (EDE) for hypothesis 2 comparing the treatment 2 with the treatment 3.

	<i>Outcome: total sum of legitimate sanctions</i>			
	(1)	(2)	(3)	(4)
Treatment	0.17*	0.07	-0.03	-0.13
	(0.08)	(0.08)	(0.08)	(0.08)
Strength of the EDE (in std. dev.)	0	0.1	0.2	0.3
Adj. R ²	0.34	0.34	0.34	0.34
Num. obs.	349	349	349	349

Notes:

Treatment tests treatment 2 against the treatment 3. The treatment 2 is a scenario where the elders did blame the town chief. The treatment 3 is a scenario where the elders did not blame the town chief. Four robustness check strategies are used. The first column is the main empirical strategy with the outcome variable being standardized. Three sensitivity analysis are used (column 2 to 4). The second, third and fourth column use an experimenter demand effect of 0.1, 0.2, and 0.3 standard deviation respectively. Cluster robust standard error at the village level. Village and block fixed effects are used in the last three columns. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

J.6 Robustness check with alternative measurement strategy: proportion of sanctions chosen

In this section we run the main analysis to test separately T1 vs. C, T4 vs. T1, and T2 vs. T3 with an alternative measurement strategy: the proportion of direct and indirect sanctions chosen as legitimate. Doing so, we can compare the relative effect size for both outcomes.

Table J.6: Robustness check for hypothesis 1 and 2 comparing treatment 1 with control, treatment 4 with treatment 1 and treatment 2 with treatment 3. Alternative outcomes are used: proportion of direct or indirect sanctions considered as legitimate.

	T1 - C		T4 - T1		T2 - T3	
	Dir. (1)	Indir. (2)	Dir. (3)	Indir. (4)	Dir. (5)	Indir. (6)
Treatment	19.22*** (1.93)	62.90*** (4.58)	4.57* (1.74)	8.92*** (2.51)	2.21 (2.10)	8.13** (3.02)
Adj. R ²	0.38	0.57	0.34	0.29	0.18	0.29
Num. obs.	374	374	361	361	349	349
N Clusters	77	77	77	77	77	77

Notes:

Treatment tests treatment 1 against the control in columns 1 and 2 with the treatment 1 being the scenario where the chiefs take a small part of the money for its own benefit. Treatment tests treatment 4 against the treatment 1 in columns 3 and 4 with the treatment 4 being the scenario where the chiefs take half of the money for its own benefit and treatment 1 where the chiefs take only a small part of the money of its own benefit. Treatment tests treatment 2 against the treatment 3 in columns 5 and 6. The treatment 2 is a scenario where the elders did blame the town chief. The treatment 3 is a scenario where the elders did not blame the town chief. Columns 1, 3, and 5 use the proportion of direct sanctions considered as legitimate. Column 2, 4, and 6 use the proportion of indirect sanctions considered as legitimate. Cluster robust standard error at the village level with village and block fixed effects. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

K Relative costs and legitimacy of sanctions

K.1 Summary statistics of the outcomes of the focus group discussions

We deployed five focus group discussions with elders of the communities in various villages of the study area. During these focus group discussions they were asked to give a costs associated to the 9 sanctions under study. The average cost is displayed for each sanction in the following table.

Table K.1: Table of the cost of sanctions

Sanction	Cost
Indirect sanctions	
Complain to the elders	2.25
Complain to the section chief	-
Complain to the paramount chief	1.25
Direct sanctions	
Blame the chief directly	1.25
Threaten the chief directly	4.5
Refuse to get married to the chief family	5
Refuse to pay local taxes	5
Refuse to work for the chief	2.25
Refuse to participate to collective labour	3.75

Notes:

The table represents the average costs of sanctions measured on a 1 to 5 scale, 5 being high cost. The measure comes from focus groups discussions undertaken by elders in 4 communities. The measure does not include complain to the section chief.

K.2 Relationships between legitimacy of the sanctions and perceived costs

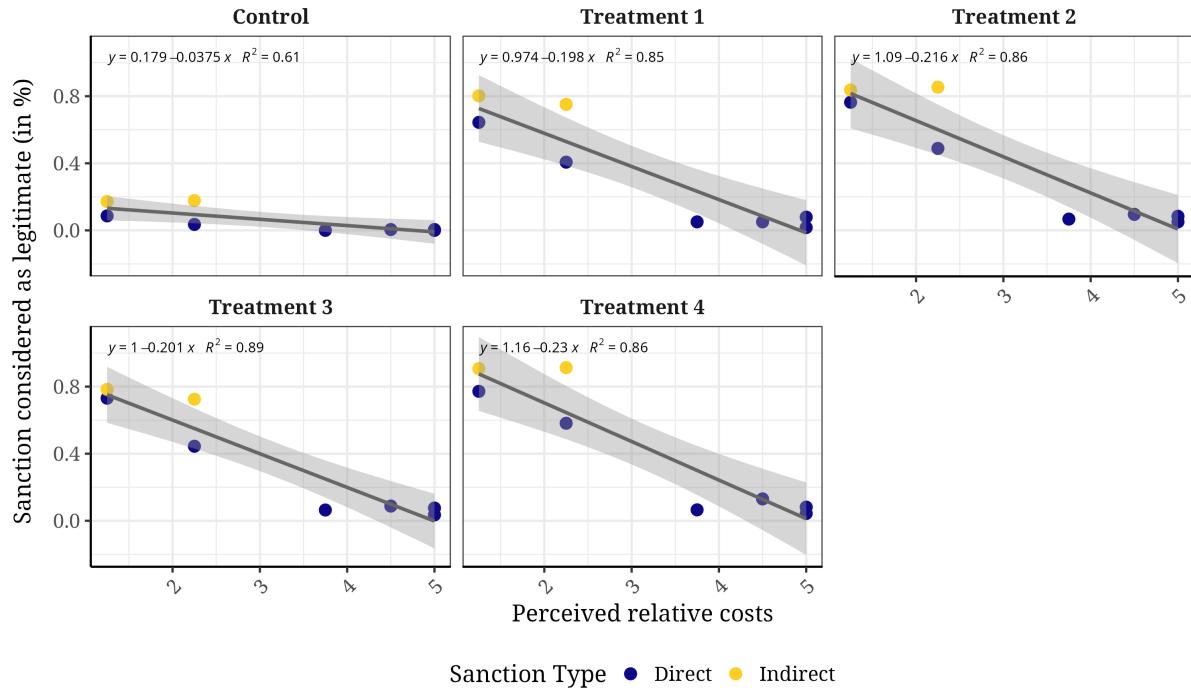


Figure K.1: Statistical association between the perceived costs of sanctions and their average legitimacy by treatment conditions.

Note: The graphic displays the statistical association between the perceived costs of sanctions and their average legitimacy across treatment conditions. Each dot represents one of the eight sanction types for which perceived cost was assessed. The fitted line and model are also shown.

K.3 Statistical test of hypothesis 1a

Table K.2: Statistical tests for the impact of sanction type on the probability of being chosen

	Outcome: sanction as legitimate (in %)	
	(1)	(2)
Indirect sanctions	0.50*** (0.06)	0.23** (0.07)
Treatment 1	0.33*** (0.09)	0.29** (0.09)
Treatment 2	0.39*** (0.10)	0.34*** (0.09)
Treatment 3	0.35*** (0.10)	0.31** (0.09)
Treatment 4	0.42*** (0.11)	0.38*** (0.10)
Costs		-0.14*** (0.02)
Observations	45	40
Adj. R ²	0.59	0.82

Notes:

The table represents the results of a statistical test for hypothesis 1a stating that less costly sanctions are less likely to be chosen. The outcome variable is the proportion of respondents considering the sanction as legitimate in each treatment arm. Indirect sanctions is a dummy being one when the sanction type is indirect. Four treatment dummies are integrated. Finally, the costs of each sanction is also integrated.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$

K.4 Additional results

Table K.3: Average treatment effects of chief malevolence on the total number of legitimate sanctions by costs of the sanctions

	T1 - C		T4 - T1	
	Low costs	High costs	Low costs	High costs
Treatment	2.84*** (0.15)	0.20*** (0.04)	0.48*** (0.12)	0.06 (0.06)
Observations	374	374	361	361
Adj. R ²	0.587	0.11	0.43	0.19

Notes:

The table represents the results of a statistical test for hypothesis 1 disaggregating the outcomes by the costs of sanctions. The outcome variable is the proportion of respondents considering the low costs sanctions as legitimate in column 1 and 3 and high costs sanctions in column 2 and 4. T1-C tests Treatment 1 where the quantity of money stolen is very small vs the control group, and T4-T1 tests Treatment 4, where half of the money is stolen vs Treatment 1 (hypothesis 1). village and block fixed effects are used. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

L Heterogeneous effect by gender, income, and voting rights

L.1 Gender

Table L.1: Heterogeneous treatment effects on the number of legitimate sanctions by gender

	General		Direct		Indirect	
	(1)	(2)	(3)	(4)	(5)	(6)
T1-C	3.07*** (0.16)	3.45*** (0.18)	1.17*** (0.09)	1.31*** (0.11)	1.90*** (0.10)	2.14*** (0.11)
T4-T1	3.69*** (0.15)	3.94*** (0.17)	1.49*** (0.09)	1.64*** (0.11)	2.19*** (0.09)	2.30*** (0.10)
Female	0.26 (0.22)		0.10 (0.11)		0.15 (0.14)	
T1-C * Female		-1.08** (0.34)		-0.38* (0.18)		-0.69** (0.22)
T4-T1 * Female		-0.68* (0.32)		-0.41* (0.20)		-0.28 (0.18)
Adj. R ²	0.60	0.61	0.42	0.43	0.61	0.62
Num. obs.	558	558	558	558	558	558

Notes: Three outcome variables are used: general, direct and indirect number of legitimate sanctions. General is the total sum of sanctions considered as legitimate. Direct is the sum of sanctions considered as legitimate targeting directly the chief. Indirect is the sum of sanctions considered as legitimate targeting an higher authority. T1-C tests Treatment 1, where the quantity of money stolen is very small vs the control group, and T4-C tests Treatment 4, where the half of the money was stolen vs Treatment 1. The models include a factor for the treatment comparison and a dummy for gender. Village and block fixed effects are used. Robust standard errors in parenthesis.
 *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

L.2 Income

Table L.2: Heterogeneous treatment effects on the number of legitimate sanctions by income

	General		Direct		Indirect	
	(1)	(2)	(3)	(4)	(5)	(6)
T1-C	3.07*** (0.16)	3.94*** (0.30)	1.17*** (0.09)	1.55*** (0.16)	1.90*** (0.10)	2.39*** (0.19)
T4-T1	3.69*** (0.15)	4.22*** (0.26)	1.49*** (0.09)	1.59*** (0.15)	2.19*** (0.09)	2.63*** (0.16)
income (Q2)	0.58+ (0.32)		0.32* (0.16)		0.25 (0.19)	
income (Q3)	0.78* (0.33)		0.39* (0.16)		0.39* (0.20)	
income (Q4)	1.21*** (0.34)		0.55*** (0.16)		0.65** (0.22)	
T1-C * income (Q2)	-0.76+ (0.45)		-0.40+ (0.24)		-0.37 (0.28)	
T4-T1 * income (Q2)	-0.34 (0.42)		0.07 (0.27)		-0.41+ (0.24)	
T1-C * income (Q3)	-1.20** (0.44)		-0.46+ (0.24)		-0.75** (0.28)	
T4-T1 * income (Q3)	-1.06** (0.39)		-0.41+ (0.23)		-0.65** (0.24)	
T1-C * income (Q4)	-1.53*** (0.45)		-0.72** (0.25)		-0.82** (0.28)	
T4-T1 * income (Q4)	-0.79+ (0.43)		-0.10 (0.25)		-0.69** (0.26)	
Adj. R ²	0.60	0.61	0.42	0.44	0.61	0.62
Num. obs.	558	552	558	552	558	552

Notes: Three outcome variables are used: general, direct and indirect number of legitimate sanctions. General is the total sum of sanctions considered as legitimate. Direct is the sum of sanctions considered as legitimate targeting directly the chief. Indirect is the sum of sanctions considered as legitimate targeting an higher authority. T1-C tests Treatment 1, where the quantity of money stolen is very small vs the control group, and T4-C tests Treatment 4, where the half of the money was stolen vs Treatment 1. The models include a factor for the treatment comparison and a factor for the income category classified into four quartiles. Village and block fixed effects are used. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

L.3 Voting rights

Table L.3: Heterogeneous treatment effects on the number of legitimate sanctions by voting right

	General		Direct		Indirect	
	(1)	(2)	(3)	(4)	(5)	(6)
T1-C	3.07*** (0.16)	2.07*** (0.33)	1.17*** (0.09)	0.56*** (0.16)	1.90*** (0.10)	1.51*** (0.21)
T4-T1	3.69*** (0.15)	2.60*** (0.31)	1.49*** (0.09)	0.80*** (0.19)	2.19*** (0.09)	1.80*** (0.20)
Voting rights		-0.75** (0.27)		-0.42** (0.13)		-0.33* (0.16)
T1-C * Voting rights			1.36*** (0.39)	0.83*** (0.20)		0.53* (0.25)
T4-T1 * Voting rights			1.43*** (0.35)	0.91*** (0.22)		0.52* (0.22)
Adj. R ²	0.60	0.61	0.42	0.44	0.61	0.61
Num. obs.	558	558	558	558	558	558

Notes: Three outcome variables are used: general, direct and indirect number of legitimate sanctions. General is the total sum of sanctions considered as legitimate. Direct is the sum of sanctions considered as legitimate targeting directly the chief. Indirect is the sum of sanctions considered as legitimate targeting an higher authority. T1-C tests Treatment 1, where the quantity of money stolen is very small vs the control group, and T4-C tests Treatment 4, where the half of the money was stolen vs Treatment 1. The models include a dummy indicating whether the respondent has the right to vote for the town chief. Robust standard errors in parenthesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$.

M Elder involvement: information and covering mechanisms

We investigate in this section why the involvement of the council of elders increase the quantity of sanctions considered as legitimate.

Two complementary mechanisms may be at play: the information-legitimation and the covering mechanism. The first proposes that when elders act against the town chief, it signals the legitimacy of villagers' grievances, encouraging them to express their concerns. The second, the covering mechanism, suggests that collective actions by villagers are seen as less isolated and thus less prone to retaliation from the town chief.

We assess these mechanisms by comparing treatment 2 and treatment 3 effects on two outcomes (see Figure M.1): a) the number of sanctions deemed legitimate (outcome 4), representing the legitimization mechanism, and b) the number of sanctions villagers feel able to take (outcome 5), representing the covering mechanism. If the effect appears only for outcome 4, it may indicate the legitimization mechanism alone.

Figure M.1 shows that the elders' influence is especially significant for indirect sanctions. Additionally, the effects on both outcomes suggest that both the legitimization and covering mechanisms are likely at play.

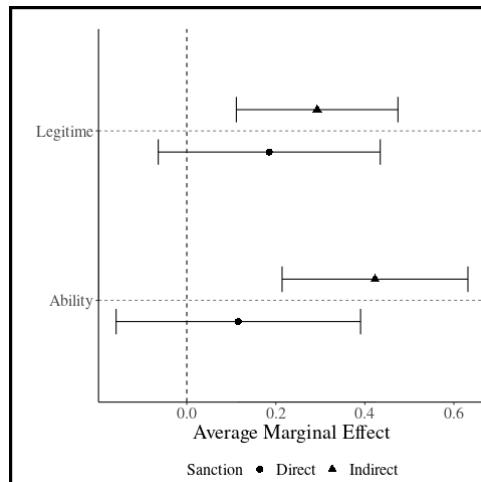


Figure M.1: Average treatment effects of elders' behavior on sanctioning preferences

Notes: The figure presents the average treatment effects of elders' behavior on the number of direct and indirect sanctions considered legitimate (top) and the number of direct and indirect sanctions the respondent would feel able to undertake (bottom). Bars represent 95% confidence intervals.

We find evidence supporting the validity of these mechanisms in select quotes from the open-ended questions, where the support of elders emerges as a crucial factor in shaping collective efforts aimed at maintaining social order. Respondents do not only condition their preferences on elders' action, but also their (stated) behaviors.

“I will report home to the elders and I will join the elders to sermoned the chief” – One respondent receiving treatment 4

“I alone will not take any action because he is our chief, but if we came as one in the village we will take the action together” – One respondent receiving treatment 4

Our research reveals that the behavior of the elders shape respondents action when confronting to chiefs malevolence. Such individuals play a central in counterbalancing power dynamics in the case of elite capture of formal institutions (Shapland et al. 2023).

N Checklist for Reporting Standards for Experiments

N.1 Hypotheses

- **What question(s) was (were) the experiment designed to address?**

How do citizens in Sierra Leone perceive the legitimacy and feasibility of sanctioning their chiefs outside of elections?

- **What are the specific hypotheses to be tested?**

Hypothesis 1: Village members view sanctions against chiefs who violate community interests as legitimate and feasible.

Hypothesis 2: community members will be less willing to view sanction as legitimate and feasible if the council of elders already blamed the chief for their action.

N.2 Subjects and Context

- **Why was this subject pool selected? Who was eligible to participate in the study?**

Communities were randomly selected from villages located near mangrove resources, with eligibility based on population size. This sampling frame was defined by the goals of the broader research project, which focused on forest-dependent livelihoods and community-forest relationships. Although this paper does not analyse forest use directly, selecting such villages ensured that local chiefs were embedded in the community, a key theoretical scope condition for this study. Villages with more than 200 households were excluded, as chiefs in larger communities are often less integrated. Villages with fewer than 20 households were also excluded due to sample size requirements in other parts of the project.

We randomly selected in each village 12 heads of household, after a full census of settlements. The requirement was that the head of household should be older than 18 years old.

- **What would result in the exclusion of a participant? Were any aspects of recruitment changed (such as the exclusion criteria) after recruitment began?**

Age below 18 years old or not being a head of the household (or a representative) would result in the exclusion of the participants.

No aspects of recruitment changed after recruitment began.

- **Procedures used to recruit and select participants. If there is a survey: Identify the survey firm used and describe how they recruit respondents.**

No survey firm was involved.

- **Recruitment dates defining the periods of recruitment and when the experiments were conducted. Also list dates of any repeated measurements as part of a follow-up.**

Data collection happened from 1st of April to the 23rd of May.

- **Settings and locations where the data were collected. In the field, lab, classroom, or some other specialized setting?**
Data was collected in the field, generally around or in the house of the participant (at its convenience).
- **Other relevant specifics of the population: e.g., large public university vs. small private university; geographic location; etc.**
The study location was the South of the country in mangrove areas.
- **If there is a survey: Provide response rate and how it was calculated.**
All participants surveyed responded.

N.3 Allocation Method

- **Details of the procedure used to generate the assignment sequence (e.g., randomization procedures).**
In each village, we block randomize on gender the assignment to one of the five treatment conditions. The survey experiment condition was then automatically loaded in the tablet.
- **If random assignment used, then details of procedure (e.g., any restrictions, blocking). Note the unit of randomization (individuals, groups, households, etc). Pay careful attention to report clustered random assignment if subjects were assigned at some level other than the individual subject.**
Individuals were randomized to one of the five treatment conditions. Village, and gender were the two blocking variables.
- **If random assignment used, provide evidence of random assignment. If demographic or other pretreatment variables were collected, a table (in text or appendix) showing baseline means and standard deviations for demographic characteristics and other pre-treatment measures by experimental group.**
Appendix E provides covariate balances.
- **If blocking was used, and group assignment proportions were not equal across blocks, provide table for each of the blocks. If there are too many blocks for this to be practical, combine blocks to present weighted averages of covariates using inverse probability weighting.**
Appendix E provides covariate balances with village and gender fixed effects.
- **Blinding: Were participants, those administering the interventions, and those assessing the outcomes unaware of condition assignments?**
Not applicable.
- **If blinding took place, include a statement regarding how it was accomplished and how the success of blinding was evaluated.**
Not applicable.

N.4 Treatments

- **Description of the interventions in each treatment condition, as well as a description of the control group.** Descriptions should be sufficient to allow replication: **Summary or paraphrasing of experimental instructions in the article text; verbatim instructions and/or other treatment materials provided in an appendix.**

The main text provides a table with description of each treatment condition.

- **How and when manipulations or interventions were administered. Method of delivery: Pen-and-paper vs. computer or internet vs. face-to-face communication vs. over the telephone.**

Enumerators read the scenario, question, and possible answers. They reported on a tablet respondents' answers. Answers were then uploaded to SurveyCTO server.

- **If computerized, the software should be described and cited. (If possible, programs should be included in appendix so as to be available for purposes of replication.)**

SurveyCTO

- **For lab experiments (and other experiments, when relevant): Report the number of repetitions of the experimental task and the group rotation protocol. Report the ordering of treatments for within-subject designs. Any piggybacking of other protocols should be reported. Report any use of experienced subjects or subjects used in more than one session or treatment. Time span: How long did each experiment last? How many sessions were subjects expected to attend? If there were multiple sessions, how much time passed between them? Total number of sessions conducted and number of subjects used in each session. Was deception used? Treatment fidelity: Evidence on whether the treatment was delivered as intended.**

Not Applicable

- **Report any instructional anomalies or inaccuracies.**

Not Applicable

- **Were subjects given quizzes on the experimental instructions?**

Not Applicable

- **Were there practice rounds? If so, how many and what were the results?**

Not Applicable

- **Did subjects complete a post-experiment debriefing, interview, or questionnaire? If so, is there evidence that subjects understood the instructions and treatments?**

Not Applicable

- **Did the experimental team observe aspects of the intervention?**

Not Applicable

- **Provide description of manipulation checks, if any. Were incentives given? If so, what were they and how were they administered.**

No incentives were given. Two questions were asked to determine whether the experimental conditions were understood by the respondent.

N.5 Results

- **Outcome Measures and Covariates** Provide precise definition of all primary and secondary measures and covariates. For indices, provide exact description of how they are formed. For survey items provide exact question wording in an appendix. Please provide a copy of the complete survey questionnaire (in an on-line appendix if it is long).

Survey instruments are provided in supplementary materials. Outcome variables index are provided in the main text, in the design section.

- **Clearly state which of the outcomes and subgroup analyses were specified prior to the experiment and which were the result of exploratory analysis.**

Pre-analysis plan and deviations from pre-analysis plan are provided in the online appendix.

- **Complete CONSORT Participant Flow Diagram**

Not applicable.

- **Researchers will conduct statistical analysis and report their results in the manner they deem appropriate. We recommend that this reporting include the following:** Report sample means and standard deviations for the outcome variables using intent-to treat (ITT) analysis (means for the entire collection of subjects assigned to a group, whether the treatment is successfully delivered or not). If the experiment uses block randomization with unequal assignment rates, present ITT analysis by block or present overall means using inverse probability weighting.

Results section and appendix provide the necessary information.

- **Note if level of analysis differs from level of randomization and estimate appropriate standard errors.**

Not applicable.

- **If there is attrition, discuss reasons for attrition and examine if attrition is related to pre-treatment variables.**

Not applicable.

- **Report for other missing data (not outcome variables): Frequency or percentages of missing data by group.**

Not applicable.

- **Methods for addressing missing data (e.g., listwise deletion, imputation methods).**

Not applicable.

- **For each primary and secondary outcome and for each subgroup, provide a summary of the number of cases deleted from each analysis and rationale for dropping the cases.**

Not applicable.

- **For survey experiments: Describe in detail any weighting procedures that are used.**

We use fixed effects by experimental block: village and gender. Results are consistent when those fixed effects are removed.

N.6 Other Information

- **Was the experiment reviewed and approved by an IRB?**

The experiment was reviewed and approved by the Office of Sierra Leone Ethics and Scientific Review Committee (SLESRC n°020/04/2023).

- **If the experimental protocol was registered, where and how can the filing be accessed?**

A Pre-Analysis Plan for this project has been registered with OSF: <https://osf.io/8r7zm>.

- **What was the source of funding? What was the role of the funders in the analysis of the experiment? Were there any restrictions or arrangements regarding what findings could be published? Any funding sources where conflict of interest might reasonably be an issue?**

We are grateful to the International Growth Center (IGC) for providing funding through their green transition scheme. No restrictions or arrangements. No conflict of interests.

- **If a replication data set is available, provide the URL.**