Cumulative knowledge in the social sciences: The case of improving voters' information*

Federica Izzo Torun Dewan Stephane Wolton

Current draft: August 22, 2018

Abstract

What happens when voters become better informed about their representatives' actions or performance? The empirical literature reports mixed findings on incumbents' electoral fortunes for both good and bad news. We introduce a political agency model, meant to inform empirical research, in which voters are fully rational, unconstrained, and unbiased. We show that researchers should not expect information campaigns to always work or produce similar effects. Identical interventions in similar contexts can give different results. The same intervention measured in different ways can give different results. The analysis also highlights that the comparability of empirical estimates is non-monotonic in the geographical spread of observations. We offer recommendations to improve the accumulation of knowledge. By constructing variables related to governance outcomes and interacting them with treatment status in their regression analyses, researchers can recover fully comparable estimates of information treatment effects. Further, to avoid attenuation bias, good and bad news should be defined in absolute, not relative terms.

Key words: Comparability of empirical estimates, pooling observations, intervention, good news, bad news

*We thank Matilde Bombardini, Peter Buisseret, Selina Hofstetter, Mark Kayser, and Janne Tukiainen for their helpful comments and advice. Corresponding author: Federica Izzo f.izzo@lse.ac.uk

1 Introduction

A policy-maker seeks to understand how improving voter information affects electoral accountability. Reviewing recent works applying gold standards of empirical research (e.g., randomized control trials, natural experiments), the policy maker's staff observes that similar information treatments yield varied effects on outcomes such as voters' knowledge of politics, politicians' behavior, or electoral results. Should our policy-maker then conclude that there is no predictable effect of increasing voters' information? Not yet, this conclusion presupposes that studies implementing similar treatments always yield estimates of the same effect; that is, these studies are necessarily comparable.

With the help of a political agency model built specifically to inform empirical work, we highlight that two studies using the same design in similar contexts (in formal language, the same model primitives) can yield two distinct true (asymptotic) effects *even after conditioning on good or bad news*. We further show that seemingly innocuous differences in interventions can generate quite distinct outcomes. In particular, results are likely to vary dramatically with the publicity of the intervention or the exact nature of the information provided, i.e., information about politicians' actions (e.g., corruption) or performance (e.g., completion of development projects).

When are studies comparable then? Take two studies implemented in two similar countries, in which after administrating the same treatment and conditioning on good and bad news, researchers simply look at the difference in average outcomes between the treated and control groups (as many randomized control trials do). If treated and control units are voters from the same village, these two studies are comparable, unless the information provided is about politicians' actions. If treated and control units are different villages, these two studies are not comparable, unless these villages are drawn from all over the country. In other words, we uncover that the comparability of identical treatments is non-monotonic in the geographical spread (level of randomization, sampling frame) of observations. Estimates are comparable for the lowest spread (treated and control come from the same village) or the highest spread (treated and control are different villages distributed all over the country). In-between (e.g., treated and control are different villages from a few geographically adjacent districts), comparability fails because, while the average treated village is always the same within (and across) sampling frames, the average control village is not.[1] Random assignment is not sufficient to keep everything constant.

What can be done to improve comparability and facilitate the accumulation of knowledge? We recommend that researchers collect data on governance outcomes (e.g., number of development projects completed) and the information to be disseminated via the intervention (e.g., account irregularities) for *all* units in the sample. Researchers should then run a fully interacted model. In this specification, a treated village is only compared with its exact equivalents in the control group (same governance outcome, same corruption level). Because of this all else equal property, fixing a level of randomization (individual or village), estimates become fully comparable across studies.

Comparability can also be improved by ensuring that researchers use consistent definitions of good and bad news. On this, researchers face a difficult task as they are never certain how voters interpret the information provided. Several approaches have been recommended to deal with this issue (e.g., benchmarking or the difference between voters' prior and information provided). We show that qualifying news on the basis of relative performance may induce attenuation bias, which does not arise when scholars instead rely on absolute performance.

We conclude this introduction by connecting our paper to the most closely related formal works (we discuss the empirical literature extensively in Sections 6 and 7). Following a long tradition (most recently reviewed in Ashworth, 2012), we use a political agency framework to model the relationship between voters and their representatives. More precisely, our work connects with a host of papers studying the consequences of providing additional information to voters (e.g., Prat, 2005; Fox, 2007; Ashworth and Shotts, 2010; Fox and Van Weelden, 2012; Ashworth and Bueno de Mesquita, 2014). However, while previous studies focus on the normative consequences of greater transparency, we are concerned with what can be learned from empirical analyses of information treatments.

As such, our work is in closest conversation with Wolton (2018) and Grossman, Michelitch, and Prato (2018). Like Wolton (2018), we evaluate what empirical estimates of information treatment effects actually measure (a new approach which has been referred to as theoretical implications

[1] This variation in outcomes in the control group arises because of politicians' strategic behavior as we discuss extensively in the main text.

of empirical models). The contexts, however, are very different. Wolton (2018) considers biased media, whereas we focus on external interventions by researchers (for RCTs) or agencies (for natural experiments). Wolton (2018), further, considers the interaction between a representative voter and a single officeholder. We, in turn, propose a theoretical framework which includes many voters, villages, districts, and states. This allows us to precisely map equilibrium outcomes into empirical quantities of interests. To the best of our knowledge, our paper is unique in this respect.

Grossman, Michelitch, and Prato (2018) present a political agency model with features analogous to ours (for example, both assume that politicians' performance is binary, and thus a coarse signal of their underlying ability or honesty). The objectives, though, are fundamentally different. Grossman, Michelitch, and Prato (2018) use their theoretical framework to generate comparative statics on the effect of performance treatments on the number of candidates and on the electoral fortune of the incumbent (a more traditional empirical implications of theoretical models approach). In turn, we establish properties of the empirical estimates of both performance and action treatments and discuss how and when these estimates are comparable. Our paper is, we believe, the first formal work to treat the possibility of accumulating knowledge from multiple empirical studies.

2 The model

Our model is a slight variation of traditional political agency models, meant to produce theoretical equivalents of empirical estimates of interest. We consider a country divided into S states (i.e., any administrative entity overarching smaller units). Each state, in turn, is composed of D districts with one representative for each. Each district is constituted of V villages inhabited by a mass 1 of voters. Slightly abusing notation, we use S, D, and V to denote both the set of units and the cardinality of this set (the number of districts could vary by state, and the number of villages by district, at the cost of complicating notation). Our set-up is flexible enough to encompass many different empirical settings. It can be adapted to the study of state legislators (in which case, $S > 1$, $D > 1$, $V \geq 1$), national legislators ($S = 1$, $D > 1$, $V \geq 1$), or village heads ($S > 1$, $D > 1$, $V = 1$). As we will see, our framework can be used to evaluate the effect of information treatments at all possible levels: across states, across districts, within districts, or even within villages.

Our model has two periods. In the first period, an incumbent (I) is in office in each district $d \in D$ of each state $s \in S$. At the end of the first period voters who live in the district sincerely cast a ballot for the incumbent or a challenger (C). In each period and each district, the office-holder exerts effort on a project (e.g., building a school) at the village level. At the end of the period, the project is either completed ($\omega = 1$) or non-completed ($\omega = 0$). The probability that the project is completed depends on two factors: (i) the representative's effort in village $v \in V$ in period $t \in \{0, 1\}$ and (ii) an underlying environment which determines the productivity of effort (in more formal terms, a state of the world). While effort is village specific and denoted e_v^t, $t \in \{1, 2\}$, $v \in V$, the underlying environment is assumed to be the same for all villages within a given state $s \in S$ and is thus denoted $\theta_s \in \{\beta, 1\}$ (with $\beta < 1$). The environment θ_s can be understood as the funds available to state legislators, the conditions of the state economy, or any other common shock which affects all the villages in the same state (notice that our approach allows for village-specific shocks, one just needs to set $S > 1$ and $D = V = 1$). We say that the environment is favorable when $\theta_s = 1$ and unfavorable otherwise. The common prior is: $Pr(\theta_s = 1) = p \in (0, 1)$, with θ_s i.i.d. across states and periods. The office-holder always learns the state before making his effort decision. For simplicity (and without much loss of generality for our implications), we assume that effort is bounded between 0 and 1 (i.e., $e_v^t \in [0, 1]$) and the probability the project is completed in a village v in state s is: $Pr(\omega_v^t = 1 | \theta_s, e_v^t) = \theta_s \times e_v^t$.

A district representative can be of one of two types: $\tau \in \{b, g\}$. A type is a politician's private information. It is common knowledge that an incumbent is a type g with probability π^I: $Pr(\tau^I = g) = \pi^I$. We assume that a random challenger is a type g with probability π^C: $Pr(\tau^C = g) = \pi^C$. When not in office, a politician gets 0. All types receive some payoff $R > 0$ from holding office (e.g., monetary and ego rents). In addition, a type-g politician also gets some benefits, which we denote α, from completing projects in the villages he represents. Thus, a bad type can be understood as a corrupt, non-congruent, or self-interested office-holder. Denote $\mathbf{e_v^t} = (e_1^t, \ldots, e_v^t, \ldots, e_V^t)$ the vector of effort in each villages. An office-holder's per period utility function assumes the following form:

$$U(\mathbf{e_v^t}; \tau) = R + \alpha(\tau) \sum_{v \in V} \omega_v^t - c(\mathbf{e_v^t}), \tag{1}$$

with $\alpha(b) = 0$ and $\alpha(g) = \alpha > 0$. To simplify the analysis, we further impose that the cost of effort is additively separable and quadratic: $c(\mathbf{e_v^t}) = \sum_{v \in V} \frac{\left(e_v^t\right)^2}{2}$.

Let us now turn to voters, who must decide at the end of period 1 whether to vote for the incumbent or the randomly drawn challenger. Each voter i cares about whether a project is completed in *her* village. In addition, her evaluation of politicians in the second period also depends on an idiosyncratic valence shock—denoted σ_i—and a shock common to all voters in the district (a district-level, state-level, or national-level shock)—denoted δ. As is common in probabilistic voting models (and as a way to simplify the analysis), we assume that σ_i is uniformly distributed over the interval $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$ and δ is uniformly distributed over the interval $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$. A voter i's first-period utility is simply ω_v^1 and her second-period utility function can be represented as:

$$u_i(\omega_v^2) = \omega_v^2 + \mathbb{I}_I(\sigma_i + \delta), \tag{2}$$

with $\mathbb{I}_I$ an indicator function equal to 1 if the incumbent I is re-elected.

We use our model to evaluate the impact of different information treatments administered before villagers cast their vote. We also study separately cases when the politician is aware of the treatment being administered and when he is not.[2] Before doing so, let us first describe the information available to voters absent any treatment; i.e., the benchmark case.

In the benchmark case, Nature determines whether a voter observes the project realization in her village $\omega_v^1 \in \{0, 1\}$ before casting her vote. We assume that there is a probability λ that each voter learns the project outcome. With the complement probability $1 - \lambda$, the voter observes nothing. This implies that in the benchmark case a proportion λ of villagers are informed about ω_v^1 (given that we assume a mass of voters). This case proceeds as follows:

0. Nature draws the states $\{\theta_s^t\}_{s \in S, t \in \{1,2\}}$ as well as politicians' types in all states and districts $\{\tau_{d,s}^J\}_{d \in D, s \in S, J \in \{I,C\}}$;

   Period 1:

1. The incumbent in each district $d \in D$ observes his type $\tau_{d,s}^I$ and the state θ_s^1 and chooses how much effort to exert: $\mathbf{e_v^1}$;

[2] There are some slight technical difficulties associated with the assumption that the incumbent is unaware of the information treatment. To circumvent these problems, one can interpret the treatment as a random event. With probability ε tending to 0, an information treatment is administrated. All our results would then carry through. We omit this additional step for ease of exposition.

2. Nature determines the outcome of the project in each village $\{\omega^1_{v,d,s}\}_{v\in V,d\in D,s\in S}$ and whether a voter in village v learns ω^1_v. All voters in village v then observe the realization of the valence shocks. They cast a vote for I or C;
3. In each district, the incumbent I is re-elected or replaced by C;

   Period 2:

4. The office-holder $J \in \{I, C\}$ in each district $d \in D$ observes his type $\tau^J_{d,s}$ and the state θ^2_s and chooses how much effort to exert: $\mathbf{e^2_v}$;
5. Nature determines the outcome of the project in each village $\{\omega^2_{v,d,s}\}_{v\in V,d\in D,s\in S}$;
6. The game ends and payoffs are realized.

We now introduce our four information treatments. For simplicity, we describe our treatments at the village level. When we turn to empirical implications (Section 4), we also discuss individual-level treatments.

- Treatment 1: voters in villages v receive information about project outcome in their village. The incumbent is not aware of the treatment before choosing his level of effort.

  Formally, in the amended timing, at stage 4, all voters in v learn ω^1_v with probability 1.
- Treatment 2: voters in villages v receive information about project outcome in their village. The incumbent is aware of the treatment before choosing his level of effort.

  Formally, in the amended timing, at stage 2, the incumbent learns that all voters in village v will observe ω^1_v and at stage 4, all voters in v learn ω^1_v with probability 1.
- Treatment 3: voters in villages v receive information about the incumbent's effort choice in their village e^1_v. The incumbent is not aware of the treatment before choosing his level of effort.

  Formally, in the amended timing, at stage 4, all voters in v learn e^1_v with probability 1 (and each observes ω^1_v with probability λ).
- Treatment 4: voters in villages v receive information about the incumbent's effort choice in their village e^1_v. The incumbent is aware of the treatment before choosing his level of effort.

  Formally, in the amended timing, at stage 2, the incumbent learns that all voters in village v will observe e^1_v and at stage 4, all voters in v learn e^1_v with probability 1 (and each observes ω^1_v with probability λ) .

Observe that treatments 1 (3) and 2 (4) differ in their timing since the incumbent is informed of the intervention only in the second case. To stress this difference, we refer to 'ex-ante intervention' when the incumbent is aware of the treatment (interventions 2 and 4), and 'ex-post intervention' otherwise (interventions 1 and 3). Interventions 1 (2) and 3 (4), in turn, vary in the nature of the information provided. We label 'performance treatment' interventions informing voters of project outcomes in their village (treatments 1 and 2). In turn, 'action treatment' refers to interventions informing voters of their representative's effort (treatments 3 and 4).

In all cases, the equilibrium concept is Perfect Bayesian Equilibrium. We further impose universal divinity to guarantee that the equilibrium is unique and none of our results are driven by multiplicity of equilibria. To limit the number of cases to be considered and simplify the analysis, but without loss of substance, we impose some parametric restrictions. We assume that (i) $\alpha\beta \geq 1$, (ii) $\psi((p + (1-p)\beta) + \frac{1}{2\phi}) < \frac{1}{2}$, and (iii) $\phi(p + (1-p)\beta) < \frac{1}{2}$. The first inequality, as we will see, implies that a type-g politician always exerts a level of effort 1 in all villages. The other two inequalities—common in probabilistic voting models—guarantee that vote shares (inequality (ii)) and winning probabilities (inequality (iii)) are always interior.

3 Equilibrium behavior

As usual, we proceed by backward induction. In the second period, the office-holder faces no electoral incentives due to our timing assumption. He simply maximizes his per-period payoff given by Equation 1. A type-b politician has no benefit from exerting effort ($\alpha(b) = 0$) and finds effort costly. Naturally defining his equilibrium strategy as a function of his type, a type-b office-holder's second-period effort is $e_v^2(b) = 0$ for all villages v in his district. In turn, a good politician values a successfully implemented project with weight $\alpha(g) = \alpha > 0$ and thus chooses $\mathbf{e_v^2}$ so as to maximize $\sum_{v \in V} \alpha Pr(\omega_v^2 = 1 | e_v, \theta_s^2) - \sum_{v \in V} \frac{(e_v)^2}{2}$. The additively separable cost and our assumption on α yield that $e_v^2(g) = 1$ in all $v \in V$.

With this, we can now turn to voters' electoral decision. A voter's choice depends on her relative evaluation of the incumbent and challenger. For the incumbent, a voter uses all the information available at the time of election. We denote $\mu_v(z, \iota_v)$ the posterior of a randomly drawn voter from village v. The first argument captures whether, independently of any information treatment, Nature informs this voter about the project outcome in her village: $z = \{\omega_v^1, \emptyset\}$ with $\omega_v^1 \in \{0, 1\}$ and $z = \emptyset$ denoting she learns nothing from Nature. The second argument captures the information treatment status in village v $\iota_v \in \{\emptyset, 1, 2, 3, 4\}$ where $\iota_v = \emptyset$ stands for the benchmark case, $\iota_v = 1$ for treatment 1 (ex-post performance treatment), $\iota_v = 2$ for treatment 2 (ex-ante performance treatment), etc. Observe that all voters have the same posterior conditional on the same information, hence we can abstain from individual subscript.

Anticipating second-period behavior, the voter expects that if the incumbent is retained, a project will be completed in her village in period 2 with probability $\mu_v(p + (1 - p)\beta)$ (ignoring arguments in the posterior function for ease of exposition). If the challenger wins, the voter expects a project is successful with probability $\pi^C(p + (1 - p)\beta)$. A voter in village v thus casts a ballot for the incumbent if and only if (with ties a zero probability event):

$$\mu_v(p + (1 - p)\beta) + \sigma_i + \delta \geq \pi^C(p + (1 - p)\beta).$$

From this, we can easily obtain aggregate vote shares given that (i) we assume throughout a unit mass of voters in each village (with λ of them informed) and (ii) the idiosyncratic shock σ_i is uniformly distributed. We obtain that the realized incumbent's vote share in village v is:

$$S_v(\iota_v) := \frac{1}{2} + \psi\big((p + (1 - p)\beta)(\lambda\mu_v(\omega_v^1, \iota_v) + (1 - \lambda)\mu_v(\emptyset, \iota_v) - \pi^C) + \delta\big),\ \iota_v \in \{\emptyset, 1, \ldots, 4\}. \quad (3)$$

Observe that our assumptions on ϕ and ψ imply that the vote share of the incumbent is always interior ($S_v(\iota_v) \in (0, 1)$) for all possible voters' posteriors. Denote $\vec{\iota} = \{\iota_1, \ldots, \iota_V\}$ the vector of information treatments in villages in district d. The incumbent's vote share at the district level is then

$$S_d(\vec{\iota}) := \frac{1}{V} \sum_{v \in V} S_v(\iota_v). \quad (4)$$

The incumbent wins re-election if and only if $S_d(\vec{\iota}) \geq \frac{1}{2}$ (ties are again a zero-probability event). There are however two dimensions to take into account. First, the incumbent is not always aware a treatment is being administrated. For information treatments 1 (ex-post outcome) and 3 (ex-post action), the incumbent believes there is no intervention. We thus denote ι_v^a the treatment status of village v as perceived by the incumbent, with $\iota_v^a = \iota$ if $\iota_v \in \{\emptyset, 2, 4\}$ and $\iota_v^a = \emptyset$ if $\iota_v \in \{1, 3\}$. The incumbent's expectation of the average vote share in his district is then $\frac{1}{V}\sum_{v \in V} S_v(\iota_v^a)$.

In addition, the realization of the valence shock δ occurs after an incumbent's first-period effort decisions. I thus treats δ as a random variable when making his effort choices. Using the assumption that δ is uniformly distributed over $[-\frac{1}{2\phi}, \frac{1}{2\phi}]$ and common to all villages in the incumbent's district, we obtain that the winning probability as perceived by the incumbent is:

$$P(I\ wins) = \frac{1}{2} + \frac{\phi}{V}(p + (1-p)\beta)\sum_{v \in V}(\lambda\mu_v(\omega_v^1, \iota_v^a) + (1-\lambda)\mu_v(\emptyset, \iota_v^a) - \pi^C). \tag{5}$$

Again our assumption on ϕ guarantees that the winning probability is interior ($P(I\ wins) \in (0,1)$) for all posteriors.

When choosing his effort level e_v^1, the incumbent considers how this will affect the voters' posterior in the village, either directly (for ex-ante action treatment) or indirectly via project outcomes (absent any perceived treatment or for ex-ante performance treatment). Importantly, Equation 5 indicates that changing voters' posterior in village v only affects the incumbent's vote share in this specific village. In addition, increasing effort in village v does not change the cost of effort in other villages. The incumbent therefore considers each village independently. Consequently, we obtain:

Lemma 1. *The incumbent's effort choice in village v depends on his type $\tau \in \{b, g\}$, the environment $\theta_s \in \{\beta, 1\}$, and the perceived treatment status $\iota_v^a \in \{\emptyset, 2, 4\}$.*

We now turn to our analysis of office-holders' first-period equilibrium behavior. We denote $e_v^1(\tau^I, \theta_s, \iota_v^a)$ an incumbent's first-period equilibrium effort choice as a function of his type $\tau^I \in \{b, g\}$, the environment $\theta_s \in \{\beta, 1\}$, and the perceived treatment status of the village $\iota_v^a \in \{\emptyset, 2, 4\}$.

First, we establish that a good politician always exerts full effort in all villages ($e_v^1 = 1$). Recall from above that a good type already exerts the highest possible level of effort absent electoral incentives. Adding the latter can only increase his willingness to put in effort.

Lemma 2. *In period 1, a good incumbent exerts effort $e_v^1(g, \theta_s, \iota_v^a) = 1$ for all $\theta_s \in \{\beta, 1\}$, $\iota_v^a \in \{\emptyset, 2, 4\}$.*

We now turn to the bad incumbent starting with the cases when the incumbent believes (correctly or incorrectly) that no treatment is taking place. From the office holder's perspective, his effort choice influences his re-election probability only via the beliefs of the λ informed voters. Such voters update positively about his type when the project in their village is completed, and negatively otherwise. Any small increase in effort thus translates into a higher probability of being re-elected. The incumbent then chooses his effort to maximize the probability of successful outcome weighted by the electoral reward from project completion. A bad incumbent exerts more effort in a favorable environment since effort is more likely to produce a successful outcome. The nature of voter updating—success and failure change voters' beliefs only if bad types are less likely to complete projects—guarantees that a bad incumbent's effort is always interior.

Lemma 3. *In period 1, in the benchmark and ex-post treatment cases ($\iota_v^a = \emptyset$), a bad incumbent in equilibrium never exerts full effort and exerts more effort when the environment is favorable ($\theta_s = 1$):*

$$0 < e_v^1(b, \beta_s, \emptyset) < e_v^1(b, 1, \emptyset) < 1$$

Suppose now that the incumbent knows that a performance treatment is being adminstrated (i.e., $\iota_v^a = 2$). The incumbent then correctly anticipates that the proportion of informed voters jumps from $\lambda < 1$ to 1. The electoral cost of a failed project thus becomes higher and pushes the incumbent to increase his level of effort. The effect is particularly strong in a favorable environment due to the higher likelihood of success.

Lemma 4. *In period 1, in the ex-ante performance treatment ($\iota_v^a = 2$), a bad incumbent in equilibrium exerts more effort than in the benchmark case for all environments: $0 < e_v^1(b, \theta_s, \emptyset) < e_v^1(b, \theta_s, 2) < 1$ for all $\theta_s \in \{\beta, 1\}$.*
The increase is higher in a favorable relative to unfavorable environment: $e_v^1(b, 1, 2) - e_v^1(b, 1, \emptyset) > e_v^1(b, \beta, 2) - e_v^1(b, \beta, \emptyset)$.

Things become slightly more involved when the incumbent knows that an action treatment is taking place. Since a good incumbent always exerts full effort ($e_v^1 = 1$), treated voters rationally anticipate that any level of effort lower than 1 can only be produced by a bad type (in formal terms, the game becomes a 'signaling game'). This has two important consequences. First, project outcomes become irrelevant for voters' electoral decision. Second, any effort strictly between 0 and 1 is a waste for a bad incumbent. It induces a cost and still lets voters learn that the incumbent is a bad type. A bad incumbent thus decides between no effort, which fully reveals his type, or full effort to mimic a good type. If the rents from office R are low, then the re-election incentives are weak and a bad incumbent is not willing to exert full effort even if doing so would significantly improve his electoral chances. In turn, if the value of holding office is sufficiently large, a bad incumbent always has incentives to exert maximum effort even if doing so only moderately improves his electoral chances. In between, a bad incumbent randomizes between the two levels of effort.

Lemma 5. *In period 1, in the ex-ante action treatment ($\iota_v^a = 4$), a bad incumbent in equilibrium exerts maximum effort ($e_v^1 = 1$) with probability η and no effort ($e_v^1 = 0$) with probability $1 - \eta$. Further, η is strictly increasing with R and there exists unique $0 < \underline{R} < \overline{R}$ such that $\eta = 0$ if $R \leq \underline{R}$ and $\eta = 1$ if $R \geq \overline{R}$.*

4 From theory to empirics

We now highlight how our results can be mapped into empirical quantities. To understand the effect of information, researchers divide a sample $\mathcal{R}$ into a treated group $\mathcal{T}$ and control group $\mathcal{C}$ (where $\mathcal{T}$ and $\mathcal{C}$ denote set and cardinality): $\mathcal{R} = \{\mathcal{T}, \mathcal{C}\}$. Assuming treatment assignment is random (randomized control trials) or as-if random (natural experiments), researchers then compare outcomes of interest in the treated and control units: for example, as we do below, the propensity to vote for the incumbent. A critical decision all researchers have to make regards the 'geographical spread' of observations. By this term, we encompass two notions: the level of randomization and the sampling frame. As we demonstrate below, different geographical spreads yield different empirical quantities with different properties.

We focus on two commonly used randomization levels: village and individual levels. For village level randomization, villages in $\mathcal{T}$ receive additional information, whereas villages in $\mathcal{C}$ correspond to the benchmark scenario (i.e., Nature reveals the outcome of the project in their village with probability λ to each villager). For simplicity, we assume that all voters (and only voters) in $v \in \mathcal{T}$ receive the treatment. This corresponds to an optimal case for researchers since there is no problem of compliance or spillover which could both produce downward bias in the estimates. For individual level randomization, first researchers draw some villages $\{v_1, \ldots, v_k\}$, $k \geq 1$, so $\mathcal{R} = \{\{\mathcal{T}_{v_j}, \mathcal{C}_{v_j}\}\}_{j \in \{1,\ldots,k\}}$, where $\mathcal{T}_{v_j}$ ($\mathcal{C}_{v_j}$) denote the treated (control) voters in village v_j, $j \in \{1, \ldots, k\}$. Then, researchers allocate individuals in village v_j to the treated—who receive additional information—and control—who only have a probability λ of learning project outcome—groups: $i \in \mathcal{T}_{v_j} \cup \mathcal{C}_{v_j} \subseteq v_j$ (slightly abusing notation we denote v_j the set of voters living in village v_j). In other terms, the randomization is blocked at the village level. While our analysis above can directly apply to village level randomization, some adaptations to our framework are required to discuss individual level randomization. We discuss these amendments below.

In turn, researchers can adopt different sampling frames. Units can be drawn from one village, from multiple villages within a district, from multiple districts within a state, or all over the country. Using our notation, when the sampling frame is individuals within a village (obviously, this is possible only for individual-level randomization), $\mathcal{R} \subseteq v$. We refer to this approach as 'within village analysis.' When the sampling frame is a district, 'within district analysis,' we have $\mathcal{R} \subseteq d$. When the sampling frame is a state, 'within state analysis,' $\mathcal{R} \subseteq s$. Finally, when the sampling frame is the whole country, 'country-wide analysis,' $\mathcal{R} \subseteq C$. Observe that in our framework we always have a mass of voters per village, villages per district, districts per state, states in the country. We can thus recover the true average treatment effect for different sampling frames.

In line with the empirical literature, we also make a distinction between treatment effect conditional on good and bad news. In our setting 'good news,' which we label n^+, is project completion ($\omega_v^1 = 1$) for performance treatments and full effort ($e_v^1 = 1$) for action treatments. In turn, 'bad news,' which we label n^-, is failed project ($\omega_v^1 = 0$) for performance treatments and non-full effort ($e_v^1 < 1$) for action treatments. We discuss problems associated with a popular alternative operationalization of good and bad news in Section 7.

Let us focus first on interventions characterised by village level randomization with a within district sampling frame. We consider the difference in incumbent's vote shares at the village level (the analysis would be similar if we look at individual voting decision within villages). Define $E_\omega(S_v(\iota_v)|\mathcal{T}, n)$ the vote share in treated villages ($\mathcal{T}$) given information treatment $\iota_v \in \{1, 2, 3, 4\}$ with news $n \in \{n^+, n^-\}$. The expected vote share for villages in the control group is $E_\omega(S_v(\emptyset)|\mathcal{C})$. Observe that we take expectations over project outcomes since voters observe success in some villages and failure in others (to the best of our knowledge, only Abida et al., forthcoming; Lierl and Holmlund, forthcoming condition on incumbent's performance for villages in the control group).[3]

[3] We take expectation over project outcome for the treated group for completeness. However, if the intervention is a performance treatment ($\iota_v \in \{1, 2\}$), project outcome is fully accounted for by n (i.e., n^+ means $\omega_v^1 = 1$ and

We denote $DTE(\iota_v, n; \tau^I, \theta_s)$ the average treatment effect for *a within district analysis*. $DTE(\cdot)$ is naturally a function of the information treatment ι_v and type of news n. Crucially, it is also a function of the incumbent's type $\tau^I \in \{b, g\}$ and environment $\theta_s \in \{\beta, 1\}$: all villages are drawn from the same district and are therefore represented by the same incumbent and under the same environment. We obtain:

$$DTE(\iota_v, n; \tau^I, \theta_s) = E_\omega(S_v(\iota_v)|\mathcal{T}, n) - E_\omega(S_v(\emptyset)|\mathcal{C}) \tag{6}$$

Let us now turn to randomization with a within state sampling frame. As above, define $E_{\omega,\tau^I}(S_v(\iota_v)|\mathcal{T}, n)$ the average vote share in treated villages ($\mathcal{T}$) given information treatment $\iota_v \in \{1, 2, 3, 4\}$ with news $n \in \{n^+, n^-\}$. The expected vote share for villages in the control group is $E_{\omega,\tau^I}(S_v(\emptyset)|\mathcal{C})$. Observe that we again take expectations over project outcome. In addition, expectations are also over incumbents' types. Villages are now drawn from multiple districts, each represented by a different incumbent. In our approach, a mass of districts are selected so all types of incumbents are included in both the control and treated groups. Consequently, the definition of the average vote share requires taking expectation over τ^I in addition to ω. We denote $STE(\iota_v, n; \theta_s)$ the average treatment effect for *a within state analysis*. In addition to the information treatment ι_v and type of news n, the STE depends on the environment $\theta_s \in \{\beta, 1\}$ since all villages belong to the same state and thus fall under the same environment. We obtain:

$$STE(\iota_v, n; \theta_s) = E_{\omega,\tau^I}(S_v(\iota_v)|\mathcal{T}, n) - E_{\omega,\tau^I}(S_v(\emptyset)|\mathcal{C}) \tag{7}$$

In turn, for village level randomization with a country-wide sampling frame, the average vote share for the treated villages is $E_{\omega,\tau^I,\theta_s}(S_v(\iota_v)|\mathcal{T}, n)$. For the control group, the expected vote share is $E_{\omega,\tau^I,\theta_s}(S_v(\emptyset)|\mathcal{C})$. The expectations for control and treated are now also over environments since units are drawn from multiple states, each with its own favorable or unfavorable environment. We denote $ATE(\iota_v, n)$ the average treatment effect for *a country-wide analysis*, which depends only, by

n^- means $\omega_v^1 = 0$). If the intervention is an action treatment ($\iota_v \in \{3, 4\}$), n is a sufficient statistic for the voter's electoral decision and thus S_v.

the reasoning above, on information treatment and the type of news. We obtain:

$$ATE(\iota_v, n; \theta_s) = E_{\omega,\tau^I,\theta_s}(S_v(\iota_v)|\mathcal{T}, n) - E_{\omega,\tau^I,\theta_s}(S_v(\emptyset)|\mathcal{C}) \tag{8}$$

We now turn to individual level randomization where voters within a village belong either to the control or treatment group. Above we have assumed that for all information treatments, all voters within a village are treated (i.e., there is no control group within a village). However, we can easily amend our approach to incorporate individual level randomization. To do so, suppose that in a village v, a proportion $\xi \in (0, 1)$ of villagers are assigned treatment $\iota_i \in \{1, 2, 3, 4\}$. The remaining $(1 - \xi)$ voters in the village learn the project outcome with probability λ, as in the benchmark $\iota_i = \emptyset$ (note that we define the treatment at the individual level, i).[4] For performance treatments the properties of the incumbent's equilibrium effort are as described in Lemmas 2-4 . For ex-ante action treatment, we will instead assume that R is sufficiently low to guarantee that a bad incumbent never has enough electoral incentives to exert full effort. He thus chooses his level of effort as if a proportion $\lambda(1 - \xi)$ of voters are informed about the village's project outcome.[5]

To investigate the effect of information treatments, we compare the propensity to vote for the incumbent among treated and controls within the village. Since we have a mass of treated and control voters, this is equivalent to looking at the incumbent's vote shares in the treated and control groups, respectively denoted with slight abuse of notation $S_\xi(\iota_i)$ and $S_{1-\xi}(\emptyset)$. Taking into account that all observations are drawn from a unique village, with a unique project, the expected vote shares are $E(S_\xi(\iota_i)|\mathcal{T}, n)$ among treated and $E(S_{1-\xi}(\emptyset)|\mathcal{C})$ among controls. We denote $VTE(\iota_v, n; \omega_v^1, \tau^I, \theta_s)$ the average treatment effect for *a within village analysis*. We obtain:

$$VTE(\iota_v, n; \omega_v^1, \tau^I, \theta_s) = E(S_\xi(\iota_i)|\mathcal{T}, n) - E(S_{1-\xi}(\emptyset)|\mathcal{C}) \tag{9}$$

In concluding this section, table 1 provides a point of reference by summarizing the discussion above. Any empirical setting can be thought as taking one (or more) combination of elements along

[4]Observe that, for ease of exposition, we assume that control and treated units are drawn from a single village. Our finding below remains the same if observations are drawn from different villages, as long as researchers include (as they often do) a village fixed effect in their regression.

[5]More generally, a bad incumbent would mix between (i) full effort ($e_1^v = 1$) and (ii) the optimal level of effort given that $\lambda(1-\xi)$ voters are informed about project outcome (and anticipate the incumbent's strategy). We assume that R is sufficiently low such that a bad incumbent plays a degenerate mixed strategy.

two dimensions: (i) the geographical spread, which encompasses the level of randomization and sampling frame (Table 1a), and (ii) the information treatment, which regroups the nature and type of information (Table 1b).

Level of randomization	Sampling frame
Village level	Within district: $DTE(\cdot)$
	Within state: $STE(\cdot)$
	Country-wide: $ATE(\cdot)$
Individual level	Within village: $VTE(\cdot)$

(a) Geographical spread

Nature	Type
Ex-post performance	Good news $\omega_v^1 = 1$
Ex-ante performance	Bad news $\omega_v^1 = 0$
Ex-post action	Good news $e_v^1 = 1$
Ex-ante action	Bad news $e_v^1 < 1$

(b) Information treatment

Table 1: Topology of empirical settings

5 The true effects of information treatments

We now discuss variations in the true treatment effects along the two dimensions described in Table 1. Proposition 1 first describes properties of the treatment effects for performance treatments. For ease of exposition, we assume away that a good incumbent in a favorable environment never produces a bad outcome (see Footnote 6 for details).

Proposition 1. *For both ex-post and ex-ante performance treatments ($\iota \in \{1, 2\}$) and for good news and bad news,*

(i) The $VTE(\iota, n; \omega_v^1, \tau^I, \theta_s)$ takes a unique value;

(ii) The $DTE(\iota, n; \tau^I, \theta_s)$ takes (a.e.) four different values, one for each $\tau^I, \theta_s \in \{b, g\} \times \{\beta, 1\}$;

(iii) The $STE(\iota, n; \theta_s)$ takes two different values, one for each $\theta_s \in \{\beta, 1\}$;

(iv) The $ATE(\iota, n)$ takes a unique value.

Proposition 1 should not be misinterpreted. Each treatment always has a *unique true* average effect for *each pair* of type τ^I and environment θ_s. However, Proposition 1 states that *across pairs* similar information treatments do not always produce the same *true* effect, even after conditioning on a particular geographical spread.[6] The extent of this problem differs across sampling frames:

[6]Proposition 1 presupposes that the researcher can draw a good type who fails in a good environment. Formally, this is equivalent to assuming that $\theta_s \in \{\beta, 1 - \epsilon\}$, with $\epsilon \to 0$. This is without much loss of generality. Point (ii) otherwise should read the $DTE(\cdot)$ takes four values for good news and three for bad news.

the variation in true treatment effects across type/environment pairs is null for within village and country-wide analyses and greatest for within district analysis.

To understand this finding, we need to go back to the definition of the various effects (Equations 6-9). For ease of exposition, we only consider good news n^+ (i.e., $\omega_v^1 = 1$). Since we condition on good news, treated units under the different sampling frames always observe the same information—a successful outcome. In each specific setting, the expected vote share for the treated group is always the same (formally, it equals $\frac{1}{2} + \psi(p + (1-p)\beta)(\mu_v(1, \iota) - \pi^C),\ \iota \in \{1, 2\}$).

The differences in treatment effects are thus due to variations in vote shares in the control group across type/environment pairs. The expected vote share in the control group depends on the proportion of informed voters who observe successful projects. For within village analyses, control and treated units are drawn from the same village. Thus, the informed voters in the control group always observe the same outcome as the treated. That is, after conditioning on good news, the proportion of informed voters in the control group who observe $\omega_v^1 = 1$ is invariant in incumbent's type and the environment.

In contrast, for within district analyses, control and treated units are drawn from different villages in the same district. Each district is characterized by a specific type and environment pair. Due to the sampling frame, in one study, researchers draw observations from a district represented by a good incumbent in a unfavorable environment; in another, from a district represented by a bad incumbent in a favorable environment, etc. Hence, in each study, observations are only a (possibly) representative sample for a unique pair: good incumbent/unfavorable environment, good incumbent/favorable environment, and so on. The researcher's analysis thus recovers the true effect only for the pair characterizing the specific pair under consideration. This is problematic for researchers since the proportion of successful projects across control villages depends on incumbent's type and environment (see Table 2). When the district from which observations are drawn is represented by a good incumbent, the proportion of successful outcomes equals $\theta_s \times 1$ (since there is a mass of villages, the proportion of successful projects equals the probability of a project being successful in a village). Similarly, a bad incumbent produces two different proportions of successful projects according to the environment. In all but knife edge cases, this yields four different proportions of successful projects, and thus four different expected vote shares for the control group. Therefore, while there is a unique treatment effect for each pair of type and environment, there are four different DTE's across pairs.

	$\tau^I = g$	$\tau^I = b$
$\theta_s = \beta$	β	$\beta \times e(b, \beta, \iota_v^a)$
$\theta_s = 1$	1	$1 \times e(b, 1, \iota_v^a)$

Table 2: Proportion of successful outcomes in the control group (within district analysis)

For within state analyses, control and treated units are drawn from different villages in different districts in the same state. Some districts are represented by a bad incumbent, others by a good one. However, all districts in the same state fall under the same environment. Hence, all observations drawn by the researcher are only a (possibly) representative sample for a unique environment, favorable or unfavorable. As the environment affects the proportion of successful projects, the expected vote share in the control group depends on θ_s. Thus, while there is a unique treatment effect for each environment, there are two different STE's across θ_s's. For country wide analyses, control and treated units are drawn across all districts and states. There is no remaining source of variation: all possible types and environments are represented in the control group, leading to a single expected vote share. Thus, by construction, the ATE is invariant in incumbent's type and environment.

Proposition 1 indicates that both the VTE and ATE are unaffected by variations in type/environment pairs. But is the VTE an unbiased estimate of the ATE? Unfortunately not, as Corollary 1 indicates. However, for ex-post information treatments, researchers can apply a simple correction to recover the ATE from the VTE. Importantly, such correction is not possible for ex-ante treatments.

Corollary 1. *For both ex-post and ex-ante performance treatments ($\iota \in \{1, 2\}$) and for good news and bad news, $VTE(\iota, n; \cdot) \neq ATE(\iota, n)$.*
For ex-post performance treatment ($\iota = 1$) and for good news and bad news,

$$VTE(1, n; \cdot) = (1 - \lambda)ATE(1, n). \tag{10}$$

As noted above, in a country-wide analysis, all outcomes, types, and environments are represented in the control group. In fact, the proportion of good and bad outcomes perfectly balance out such that, in expectation, the incumbent's vote share is as if all voters were uninformed about project outcomes (this follows from the vote share being linear in the posterior, see Equation 4, and the martingale property of posterior). In turn, the control group for a within village analysis is composed of a proportion $(1-\lambda)$ of uninformed voters and λ of informed voters. Since these informed voters have the same information as the treated, they react the same way. The effect is thus entirely driven by the proportion $(1-\lambda)$ of uninformed voters. As such, the correction we propose in Equation 10 can be thought of as dividing the VTE by the proportion of compliers among the control group.

Let us stress again that this correction does not carry through for ex-ante performance treatment. The reason is that a bad incumbent adjusts his effort to the proportion of informed voters (see Lemma 4). Since a greater proportion of voters are treated in a country-wide compared to a within village analysis, a bad incumbent's levels of efforts are different in the two settings. Hence, simply adjusting for the compliers is not enough to recover the ATE.

What can we say about the relative magnitude of the effects for the different sampling frames and type-environment pairs? The size of treatment effects depends on the proportion of informed voters in the control group who observe success. The higher this proportion, the greater the incumbent's vote share in the control group and thus the lower the impact of good news and the bigger (in absolute value) the effect of bad news. Using Table 2, the estimate conditional on good news is thus smallest for within district analysis for a good type in a favorable environment and largest for a bad type in an unfavorable environment. The estimates from within state analysis fall somewhat between these two extremes since the control group is composed of districts represented by good and bad incumbents.

Corollary 2. *For all $\iota \in \{1,2\}$ and all $\omega_v^1, \tau^I, \theta_s \in \{0,1\} \times \{b,g\} \times \{\beta,1\}$, $DTE(\iota, n; g, 1) < STE(\iota, n; 1) < STE(\iota, n; \beta) < DTE(\iota, n; b, \beta)$ for all $n \in \{n^-, n^+\}$.*

The next result shows that, at least for ex-post performance treatment ($\iota = 1$), the sign of the true effect is as expected. For such information treatment, good news always improves the incumbent's vote share, whereas bad news always decreases it.

Corollary 3. *For ex-post performance treatment ($\iota = 1$) and all $\omega_v^1, \tau^I, \theta_s \in \{0,1\} \times \{b,g\} \times \{\beta,1\}$,*
(i) $VTE(1, n^-; \omega_v^1, \tau^I, \theta_s) < 0 < VTE(1, n^+; \omega_v^1, \tau^I, \theta_s)$;

(ii) $DTE(1, n^-; \tau^I, \theta_s) < 0 < DTE(1, n^+; \tau^I, \theta_s)$*;*

(iii) $STE(1, n^-; \theta_s) < 0 < STE(1, n^+; \theta_s)$*;*

(iv) $ATE(1, n^-) < 0 < ATE(1, n^+)$*.*

Similar results do not necessarily hold for ex-ante performance treatment. Two opposite forces explain this result. To understand them, consider good news (a similar reasoning holds for bad news). Due to the intervention, more villagers are informed about the project outcome in treated than control villages. This tends to raise the incumbent's vote share. Treated voters, however, anticipate that a bad incumbent exerts more effort as a result of the treatment (Lemma 4). As a consequence, a successful project outcome becomes less informative of an office-holder's type. Conditional on success, treated voters therefore hold lower evaluation of the office holder than informed voters in the control group. This tends to decrease the incumbent's vote share. Depending on the relative strength of the extensive (number of informed voters about outcome) and intensive (relative evaluation of the incumbent in treated and control groups) margins, the treatment effect can be positive or negative.

Having discussed magnitude and sign of treatment effects, we now consider how the true effects change when all villagers learn the project outcome in their village via Nature (i.e., $\lambda \to 1$). In this scenario, all villagers in both the treated and control groups obtain information about the incumbent's accomplishment in their village. Yet, performance treatments always yield a null effect only when researchers use a within village sampling frame.

Proposition 2. *For all* $\iota \in \{1, 2\}$*, as all voters become informed by Nature (*$\lambda \to 1$*), the true effect is null for all* $\omega_v^1, \tau^I, \theta_s \in \{0, 1\} \times \{b, g\} \times \{\beta, 1\}$ *and all news* $n \in \{n^+, n^-\}$ *only for within village analysis.*

This proposition follows again from the composition of the control group. For good news (a similar reasoning holds for bad news), the effect of performance treatment is null if and only if all villagers in the control group observe project success as the treated villagers do. This condition is always met in within village analysis, since all observations are drawn from the same village, but it does not generally hold for other sampling frames. There, for all but one pair of type/environment (good incumbent/favorable environment), some villages in the control group experience a failed outcome. These villages drive the positive effect of good news in within district, within state, and country wide analyses. As we discuss in the next section, Proposition 2 indicates that empirical studies which simply compare vote shares in treated and control groups (as most works do) capture more than the impact of informing voters.

We now turn to action treatments. As discussed in the previous section, action treatments can change the nature of the game being played. Nonetheless, true effects still vary with type and environment as Proposition 3 highlights.

Proposition 3. *For both ex-post and ex-ante action treatments ($\iota \in \{3,4\}$) and for good news and bad news (when available),*
(i) The $VTE(\iota, n; \omega_v^1, \tau^I, \theta_s)$ takes two different values, one for each $\omega_v^1 \in \{0,1\}$;
(ii) The $DTE(\iota, n; \tau^I, \theta_s)$ takes at least two different values, at least one for each $\theta_s \in \{\beta, 1\}$;
(iii) The $STE(\iota, n; \theta_s)$ takes two different values, one for each $\theta_s \in \{\beta, 1\}$;
(iv) The $ATE(\iota, n)$ takes a unique value.

While similar in spirit, the findings for performance and action treatments differ when it comes to within village and within district analyses. In performance treatments, the treated voters and the informed voters in the control group base their electoral decision on the same element: success or failure of the village project. Not so much for action treatments. Treated voters consider the incumbent's effort, whereas informed voters in the control group still use project outcome.

For within village analyses, one study draws observations from a village with successful project, the other from a village with failed project. Hence, observations are only a (possibly) representative sample for a unique project outcome: failure or success. Using the same reasoning as above, this implies that while there is a unique treatment effect for a specific ω_v^1, there are two different VTE's across outcomes.

For within district analyses, the difference between performance and action treatments comes from the inference treated voters draw from the information provided. In performance treatments, both a good and a bad type can produce a failed outcome. Conversely, in action treatments bad news ($e_v^1 < 1$) can only come from a bad type since a good incumbent always exerts maximum effort (e.g., a honest office-holder never engages in corruption).[7] Hence, the only source of variation after controlling for bad news is the environment. In one study, researchers draw observations from a

[7] It should also be noted that in ex-ante action treatment, bad news are always associated with failed project since a bad type exerts zero effort (Lemma 5). This is, however, an artefact of our project production technology.

district represented by a bad incumbent in a favorable environment, in another, from a district still represented by a bad incumbent, but under an unfavorable environment. The expected vote shares in the control group in the two studies are different. Hence, as above, there are two different DTE's across environments. While only bad incumbent produces bad news, good news can be generated by both good and bad types (i.e., for ex-ante action treatment and sufficiently large office rents, see Lemma 5). Each type/environment pair yields a different expected vote share in the control group. The $DTE(\cdot)$ then takes four different values, one for each pair.

The strategic reaction of a bad incumbent to an ex-ante action treatment can also have significant consequences for the signs of the true effects. For sufficiently large office rents R, a bad incumbent imitates a good type with very high probability. This implies that in treated villages, as R increases, good news becomes less and less of an informative signal. In the extreme, when $R \geq \overline{R}$ and a bad type always exerts full effort, good news is completely uninformative about type. In this case, the expected vote share in the treated group is as if voters had received no information. In contrast, in villages in the control group a bad incumbent always exerts an intermediate level of effort. Irrespective of the value of R, observing a successful project is therefore still a positive signal about the incumbent's type. For R large enough, the treated voter's evaluation of their representative conditional on good news is thus lower than the evaluation of the informed voters in the control group conditional on project completion. If the share of successful projects in the control group is sufficiently large (e.g., observations are drawn from a district represented by a good type in a favorable environment), ex-ante action treatment can yield a *negative* true effect even after conditioning on good news. Observe that the incumbent's strategic reaction to the treatment is key. In ex-post action treatment, only good type exerts full effort and produces good news. Hence, treated villagers hold higher evaluation of their representative than voters in the control group and the impact of good news is always positive.

Corollary 4. *For bad news, the true treatment effect is always negative.*

For good news, the true effect is always positive when the following conditions are not met:

1. *The analysis is within district or within state,*
2. *The action treatment is ex-ante ($\iota = 4$), and*
3. *The rents from office are above a threshold $R^+ \in (\underline{R}, \overline{R})$.*

When conditions 1-3 are met, the true treatment effect of good news can be negative.

So far, we have focused on performance and action treatments separately. Corollaries 3 and 4 already hint that researchers should expect the nature of information to yield different true effects. Information about performance need not have the same impact as information about action. Corollary 5 establishes that this holds even after fixing the sampling frame, the type of news, project outcome, incumbent's type, and environment.

Corollary 5. *Take any two treatments $\iota \in \{1, 2\}$, $\iota' \in \{3, 4\}$, then (a.e.):*
(i) $VTE(\iota, n; \omega_v^1, \tau^I, \theta_s) \neq VTE(\iota', n; \omega_v^1, \tau^I, \theta_s)$;
(ii) $DTE(\iota, n; \tau^I, \theta_s) \neq DTE(\iota', n; \tau^I, \theta_s)$;
(iii) $STE(\iota, n; \theta_s) \neq STE(\iota', n; \theta_s)$;
(iv) $ATE(\iota, n) \neq ATE(\iota', n)$.

Voters draw different inferences when they receive information about the incumbent's action and his performance. Good news about effort do not mean the same as good news about project outcomes. Indeed, information about action is a sufficient statistic for updating about the incumbent's type. For each sampling frame, performance and action treatments will therefore yield different true effects. This is true even in the case of ex-post treatments where the incumbent does not adjust his effort choice.

6 Learning from multiple studies, problems and recommendations

In recent years, multiple empirical studies have analyzed the effect of information on incumbents' electoral performances. These works, according to the literature, yield mixed results. "Despite their prevalence, we have little hard evidence that voter information campaigns work in practice. Moreover, the evidence that does exist paints a mixed picture" (Dunning et al., forthcoming, page 8). "This research also addresses mixed findings from previous studies on the impact of information on voter behaviour" (Gottlieb, 2016, page 144). "[A] wealth of recent studies examining the effects of informational campaigns on electoral accountability and collective action yield mixed findings (e.g. Banerjee et al. 2010,2011; Björkman Nyqvist, de Walque and Svensson 2017; Chong et al. 2015; de Figueiredo, Hidalgo and Kasahara 2013; Dunning et al. forthcoming; Ferraz and Finan

2008; Humphreys and Weinstein 2012; Larreguy, Marshall and Snyder 2018; Lieberman, Posner and Tsai 2014; Olken 2007)" (Bhandari, Larreguy, Marshall, 2018, page 2).

Various explanations have been advanced as to why information campaigns may not work. One possibility is that voters struggle to absorb the information provided. Alternatively, voters face collective action problems that prevent them from using the new information efficiently. Or voters do not base their decision on the information provided, but on co-ethnicity, clientelistic ties, etc. Dunning et al. (Forthcoming, page 9) provide a comprehensive overview of the existing rationales in the literature.

The implicit assumption in all these explanations is that if voters are fully rational, unconstrained, and unbiased, then information campaigns *should* always work and produce similar effects. This assumption, our analysis shows, is fundamentally flawed. In our theoretical framework, voters can perfectly process information, face no collective action problem, and information always plays a role in their electoral decision. Further, we consider an ideal setting in which all villages, districts and states are ex-ante identical. Yet, depending on the geographical spread and nature of information, information treatments can have very different effects that vary in magnitudes and also (sometimes) in signs even after controlling for good or bad news.

Importantly, for each randomly drawn representative sample, there always exists a unique true treatment effect. Our results are therefore not driven by multiple equilibria. Our claims are also unrelated to sampling variability, and thus completely distinct from recent discussions on the use of RCTs (e.g., Deaton and Cartwright, forthcoming). The issue we document is more fundamental. The most carefully executed empirical study may only draw a representative sample for a particular environment, a particular type, or a particular combination of both depending on the geographical spread of observations. The effect of information depends on the particular representative sample being analyzed. As the previous section establishes, within district analysis can generate up to four true effects, one for each pair of incumbent's type and environment; within state analysis up to two true effects, one for each environment.

Can our explanation shed some light on the mixed findings in the literature? Twelve of the studies cited in Bhandari, Larreguy, Marshall (2018) are relevant for us (we count the 6 Metaketa field experiments in Dunning et al., forthcoming, and exclude Björkman Nyqvist, de Walque and

Svensson (2017), Lieberman, Posner and Tsai (2014), Olken (2007) which do not look at electoral outcomes).

Of these, three randomize at the individual level (Buntaine et al., forthcoming; Lierl and Holmlund, forthcoming; Boas, Hidalgo, Melo, forthcoming) and seven perform what we refer to as village-level randomization (i.e., randomization at the polling station, precinct, village, or municipality), with Humphrey and Weinstein (2012) and Platas and Raffler (forthcoming) doing both. Further, the geographical spread of observations in these works varies significantly from politicians within a single municipality (de Figueiredo, Hidalgo, Kasahara, 2013) to country-wide analysis (Ferraz and Finan, 2008; Humphreys and Weinstein, 2012; Larreguy, Marshall, and Snyder, 2018; Adida et al., forthcoming; Buntaine et al., forthcoming; Lierl and Holmlund, forthcoming). The other studies are somewhat in between within district and within state analysis. For example, Banerjee et al. (2011) consider vote shares in polling stations for 10 legislators from Delhi's state legislature. Mapping this spread into our model, this corresponds to a within state analysis with one state (Delhi) and 10 districts. In turn, Platas and Raffler (forthcoming) can be thought substantively as a form of within district analysis. The authors draw observations from four constituencies in close geographic proximity so all units arguably fall under the same environment and the small number of representatives make it unlikely all types are represented in the sample. Hence, even before accounting for the difference in treatments (action/performance, ex-ante/ex-post) and specifications, or allowing for sampling variability, our theory suggests that there are strong reasons to expect variation in results across these empirical studies.

Our framework, more generally, highlights that estimates across empirical works are comparable only when two conditions are met. First, it is necessary that all studies administer the same treatment according to our classification (Table 1b). For example, Buntaine et al. (forthcoming) provide information about embezzlement, which in our framework can be interpreted as an action treatment. In turn, Lierl and Holmlund (forthcoming) provide information about, among other things, whether the village primary school has functioning latrines, which corresponds to a performance treatment in our setting. As Corollary 5 highlights, the two studies are unlikely to yield comparable results even though the geographical spread is the same (individual level of randomization, country-wide analysis with village fixed effects).

Once this first condition is met, researchers also need to ensure that the geographical spread across studies is the same *and* allows for comparability. Propositions 1 and 3 show that for randomization at the village level (some villages are treated, others villages are in the control group), only country wide analyses can yield unbiased estimates of a unique true effect.[8] Studies which use individual level randomization with individual level outcomes are also comparable whenever they administer a performance treatment. For action treatments there are instead two true effects, which vary with the office-holder's project performance in the village.

The paragraph above highlights a non-monotonicity in comparability. Within village (in some context) and country-wide analyses always yield a single true effect. In these cases, differences in the estimates must be attributed to sample variability or to the frictions previously identified in the literature (irrationality, collective action problem, preference bias). Conversely, for within district (or close equivalent) and within state analyses, it is hard for researchers to establish whether mixed findings are due to such noises and frictions or simply reflect the variation in true effects.

So far, we have focused on comparison across studies, but our framework also speaks to comparison within a single study. Researchers sometimes are interested in validating their individual level outcomes using between village comparison (e.g., Adida et al., forthcoming; Larreguy et al., forthcoming). Our framework suggests that scholars should be careful in interpreting discrepancies in findings across different levels of aggregation. In fact, because changing the sampling frames changes the control group, the obtained estimates can differ significantly in term of magnitude. It is difficult to predict whether the effect is larger at the individual or at the aggregate levels. Further, for ex-ante action treatment coupled with bad news, the estimate changes signs whenever the treated village experiences a successful project while most of the villages in the control group do not.

Remark 1. *Suppose a proportion ξ of voters in a village v receive treatment $\iota \in \{1, 2, 3, 4\}$. For such intervention, the difference between the associated $VTE(\cdot)$ and $DTE(\cdot)$ is ambiguous.*

Further, for ex-ante action treatment ($\iota = 4$) and bad news ($n = n^-$), there exists a non-measure zero

[8] Another way to recover an estimate of the ATE is pooling data from multiple within state or within district studies (e.g., the Metaketa initiative). One caveat though is that all pooled works should administer the same treatment (Corollary 5). Otherwise, the meta-analysis only provides an estimate of the average of two or more ATE's and is no longer comparable with other country-wide analyses.

set of parameter values such that with strictly positive probability, sampled villages yield $VTE(\cdot) < 0$ *and* $DTE(\cdot) > 0$.

Remark 1 also speaks to the debate on the usefulness of conducting pilot studies. Researchers have defended localised interventions in the hope that "they will help the public sector develop similarly-successful large-scale versions of these programs" (Davis et al., 2017, page 2). A well-known problem of scale-up is that they can induce strategic responses from representatives that transform an ex-post intervention into an ex-ante one (e.g., Acemoglu, 2010). Our theoretical findings suggest another issue with the use of pilots even absent any change in office-holders' behavior: results vary with the geographical spread. Large-scale versions of successful small-scale interventions can yield disappointing result. As such, our framework may explain why Humphreys and Weinstein (2012, for electoral outcomes) and Grossman, Humphreys, Sacramone-Lutz (2014 and 2018, for politicians' effort), find significant estimates only for their within village analysis.

The discussion above should not be misinterpreted. It does not claim that the accumulation of knowledge is impossible when it comes to information treatments. Indeed we have shown that, for country-wide or within village analysis (with performance treatments), studies adminstrating the same treatment do yield comparable estimates (i.e., estimates of the same true effect). A first simple and costless step to improve comparability is then for future works to document the geographical spread of their observations and the information treatment they administer along the lines of Table 1. This would permit to understand whether variability in results is due to noise, frictions, or the estimates may simply refer to different true effects.

Further, our analysis so far has been restricted to the commonly used difference in means approach. Other empirical specifications may be more favorable to cumulative knowledge. Indeed, we now propose an alternative approach which ensures that for a given level of randomization, identical treatments always yield comparable estimates irrespective of the sampling frame.

Our recommendation is that for *all* villages in their sample, researchers collect data regarding (i) what voters can observe about politician's performance absent treatment (i.e., in our model, data about whether projects are successful in the village) and (ii) the pieces of information they would like to disseminate (e.g., malfeasance). Researchers should then run a fully interacted model including the type of news, the incumbent's performance in the village (treated or not), and the administrated treatment.

To the best of our knowledge, scholars do not generally run the specification prescribed above. Of the papers we surveyed, only Abida et al. (forthcoming) and Lierl and Holmlund (forthcoming) follow this approach. Using our framework as illustration, we now detail why a fully interacted model solves most of the issues we document. Let Y_{vst} be the incumbent vote's share in village v. Denote $P_{vds} = 1$ if the project was successful in village v in district d in state s, and $P_{vds} = 0$, otherwise. Denote $N_{vds} = 1$ for good news (i.e., success for performance treatment, full effort for action treatment). Finally, denote $T_{vds} = 1$ if village v is treated. The full interaction model we recommend running is then, respectively for performance treatment and action treatment:

$$Y_{vds} = a_0 + a_1 P_{vst} + a_2 T_{vst} + a_3 P_{vst} \times T_{vst} + \epsilon_{vst} \tag{11}$$

$$Y_{vds} = b_0 + b_1 P_{vst} + b_2 T_{vst} + b_3 N_{vst} + b_4 P_{vst} \times T_{vst} + b_5 T_{vst} \times N_{vst} + b_6 P_{vst} \times T_{vst} \times N_{vst} + \epsilon_{vst} \tag{12}$$

Two observations are worth making. First, for performance treatment (Equation 11) we do not need to include an interaction between project outcome and good news since performance *is* the news (i.e. $P_{vst} = N_{vst}$). Second, for action treatment (Equation 12) it is not sufficient to control and interact for good news. Researchers must also control and interact for observable performance (project outcome in our setting).

The estimates obtained via Equation 11 can easily be interpreted. For performance treatments, the effect of providing good news is $\widehat{a_2} + \widehat{a_3}$, the effect of providing bad news is $\widehat{a_2}$. For action treatments, the effect depends on the project outcome voters would have observed absent the treatment. For example, the effect of providing good news (i.e., full effort) is $\widehat{b_2} + \widehat{b_5}$ when voters would have observed project failure, and $\widehat{b_2} + \widehat{b_5} + \widehat{b_6}$ in cases when voters would have observed success. The resulting estimates can also easily be mapped back into our framework. For example, our analysis predicts that $\widehat{b_2} + \widehat{b_5}$ is always positive, whereas $\widehat{b_2} + \widehat{b_5} + \widehat{b_6}$ should be negative for ex-ante treatments in settings with high rents from office (see Corollary 4).

Observe that the empirical estimates obtained via Equation 11 and Equation 12 measure very different quantities than the ones obtained without adding the controls we prescribe. The two approaches, in fact, rely on two different baselines. In the commonly used difference in means, the baseline is the incumbent's vote share for the average outcome performance in the sample absent treatment. The treatment thus combines the effect of informing voters and the effect of the incumbent performing above or below average (for good and bad news, respectively). The issue we highlight is that average performance is not constant across and within sampling frames. For example, for within district analyses, the average performance is a function of both the incumbent's type and the environment. Hence, there exist multiple true effects. In contrast, in the specifications we recommend, the baseline is the incumbent's vote share absent treatment everything else constant. A treated village is therefore only compared with its exact equivalents in the control group (e.g., for action treatment, a village treated by good news which experiences project completion is only compared with control villages in which the project is successful). The specifications we suggest control for what the treated group would have observed absent treatment and for what the control group would have observed if treated. Hence, the only remaining difference between the two groups is the effect of informing voters.

The following scenario usefully illustrates the difference between the two approaches. Suppose that absent any treatment all villagers observe the project outcome in their village (formally, $\lambda \to 1$). Assume further that a researcher administers a performance treatment. Proposition 2 indicates that the common difference in means without controls or interaction should yield a non-zero effect (even after conditioning on good and bad news). Conversely, estimates obtained from Equation 11 and Equation 12 are predicted to equal zero since these specifications only measure the effect of providing information to voters.

The main advantage of the approach we propose is that for a given level of randomization, all identical interventions are fully comparable (no matter the sampling frame). That is, fixing an information treatment (Table 1b), researchers using Equation 11 or Equation 12 always recover estimates of the same unique true effect of providing specific information, whether observations are drawn from a single district, several districts within a state, multiple states, or across the country. Indeed, the specification we prescribe eliminates all variation in true estimates due to different proportions of successful projects in the control group. The remaining variations in the empirical results can only be due to statistical noise or frictions.

7 Good news, bad news, and null findings

So far, we have highlighted how our framework can explain why different studies find effects of different magnitudes, some large, some small. The signs of the predicted estimates are conform with intuition, at least for ex-post treatments. Good news increases the incumbent's expected vote share, and bad news decreases it. However, several recent studies have found that information treatments have little effect. For example, in concluding the Metaketa I, the principal investigators, Dunning et al. remark that there is "no evidence of impact of the common informational intervention across all studies in the aggregate, and little evidence of substantial impact in any of the individual studies" (forthcoming, 452). One obvious explanation discussed in the literature is that "the treatments [are] neither strong enough nor salient enough to affect voters' beliefs about politicians." (ibid., 453). We do not deny the importance of these factors, which can easily be accommodated in our model. Here, we offer an alternative explanation which has to do with the operationalization of good and bad news.

Several authors have argued that scholars should use benchmarked performance to define good and bad news (e.g., Humphreys and Weinstein, 2012; Gottlieb, 2016). This approach is meant to solve a fundamental problem for researchers: they are not always certain how voters interpret the information provided. For example, Humphreys and Weinstein (2012, 23, emphasis added) explain that "[t]he difficulty with assessing the effects of information on voter attitudes is that whether or not the information is new and *whether it is good news or bad news* depends on both the prior attitudes of respondents and the characteristics of politicians." Benchmarking at first sight seems an obvious way to address this issue. Higher relative performance can be thought as a positive signal since a piece of news is good if it is more likely to be generated by a good type.

The problem with this strategy is that it relies on the assumption that relative performance is monotonic in politicians' type. A good office-holder's performance is always higher than a bad one's. This assumption, however, needs not always be true. To see that, consider ex-post performance treatments and suppose that the productivity of effort in an unfavorable environment (β) is low. In this case, bad politicians in a favorable state generate more successful projects than good politicians in an unfavorable state: $e(b, 1, \emptyset) > \beta$. Hence, informing voters that their incumbent is better than the median politician can be essentially uninformative. More problematically, under some conditions, voters would correctly infer that their incumbent is a bad type if he is only somewhat better than the median office-holder, or a good type if he is only somewhat worse.[9]

Another strategy to define good and bad news is to compare relative performance with voters' prior expectations (e.g., Dunning et al., 2018). This approach, however, is likely to generate attenuation bias *even if* relative performance is monotonic in politicians' types. To illustrate this problem, suppose (slightly modifying our model) that in each village, there are four equally sized groups of voters each with a different prior belief on the incumbent's type. Group 1 villagers have a lower prior (π_1^I) than group 2 ($\pi_2^I > \pi_1^I$), group 2 than group 3 ($\pi_3^I > \pi_2^I$), group 3 than group 4 ($\pi_4^I > \pi_3^I$), with $\sum_{k=1}^{1} \pi_k^I/4 = \pi^I$. These different priors then map into different expectations of performance, with group 1 and group 2 voters expecting worse performance than average, and group 1 worse performance than group 2, etc. (average expected project completion is: $\pi^I(p + (1-p)\beta) + (1-\pi^I)(pe(b, 1, \cdot) + (1-p)\beta e(b, \beta, \cdot))$ increasing in π^I). We then denote $P_{ij} = k$ if voter i's expectation of incumbent j's performance belongs to the k group, or equivalently to the kth quartile. In turn, let us suppose that actual performance is such that the first quartile contains bad incumbents in an unfavorable environment, the second bad incumbents in a favorable environment, the third quartile good incumbents in an unfavorable environment, and the top quartile good incumbents in a favorable environment. We then denote $Q_j = l \in \{1, 2, 3, 4\}$ if incumbent j's actual performance belongs to the lth quartile.

Following the literature, information (Q_j) that is weakly better than voter i's prior (P_{ij}) is defined as good news. This definition wrongly supposes that all voters such that $Q_j - P_{ij} \geq 0$ are treated with information favorable to the incumbent. Take a difference of 1: $Q_j - P_{ij} = 1$ (similar reasoning holds for $Q_j - P_{ij} = 0$ or $Q_j - P_{ij} = -1$). This groups together voters in the first quartile ($P_{ij} = 1$) learning that their incumbent is in the second quartile ($Q_j = 2$), voters in the second quartile ($P_{ij} = 2$) learning that their incumbent is in the third quartile ($Q_j = 3$), and voters in the third quartile ($P_{ij} = 3$) learning that their incumbent is in the top quartile ($Q_j = 4$). However, these three groups of voters do not make the same inference. Better than average performance ($Q_j = 3$ or $Q_j = 4$) fully reveals that the incumbent is a good type; worse than average ($Q_j = 1$ or $Q_j = 2$) that the incumbent is a bad type. After learning Q_j, the posteriors of voters in the second

[9] For example, suppose that $\pi^I = p = 1/2$ and β is such that $e(b, 1, \emptyset) > \beta$ holds. Then the quartile of top performers is constituted of good types in a favorable environment, the second quartile of *bad* incumbents in a favorable environment, the third of good types in an unfavorable environment.

and third quartile thus increase (to 1); the posteriors of voters in the bottom quartile drop (to 0). That is, the set of observations such that $Q_j - P_{ij} = 1$ contains cases treated with actual good news (voters such that $P_{ij} = 2$ and $P_{ij} = 3$) and cases treated with actual bad news (voters such that $P_{ij} = 1$), which generates an attenuation bias.[10]

Overall, this section highlights that the choice of a definition of good and bad news will have important consequences down the line on the magnitude of the estimated effect of information. In particular, studies operationalizing good news as a positive difference between information and prior tend to produce downwardly biased estimates.[11] To avoid this issue, when information can be meaningfully reduced to binary categories (corruption or not, project completed or not in the village) good and bad news should be defined in absolute terms (as we do in Equation 11 and Equation 12), rather than using a benchmark, priors, or both. When information cannot easily be divided into good and bad news, researchers should adopt a more agnostic approach (e.g., as in Ferraz and Finan, 2008). For example, suppose the number of successful projects in a village varies from 0 to $z > 1$: $X \in \{0, \cdots, z\}$. Researchers should adapt Equation 11 and run a fully interacted model with dummy variables for each value of X. In this case, if voters correctly process information, the baseline effect (no project completed) should be negative and the coefficient on the interaction of treatment with the highest number of project ($X = z$) should be positive. Accounting for the possibility of non-monotonicity, the researcher should not have expectations for the sign and magnitude of the estimates of the other interacted terms.

8 Conclusion

Are empirical studies administrating identical treatment in similar contexts always comparable? Not necessarily when it comes to information treatment, our paper suggests. When scholars use the standard difference in means approach, the same interventions can yield one, two, or more

[10]This result, it should be noted, is not driven by the binary types. It would continue to hold when types are continuous. The key assumption is that villagers do not observe the environment. Under some conditions on the distribution of types, voters then correctly attribute a performance better than their prior expectation to a worse than average incumbent working in a favorable environment.

[11]Notice that researchers are always more likely to find a statistically significant effect on voters' beliefs than the incumbent's vote share. Recall that the vote share in a village is $S_v(\iota_v) := \frac{1}{2} + \psi\big((p + (1-p)\beta)(\lambda\mu_v(\omega_v^1, \iota_v) + (1-\lambda)\mu_v(\emptyset, \iota_v) - \pi^C) + \delta\big),\ \iota_v \in \{\emptyset, 1, \ldots, 4\}$ (Equation 3). Hence, the effect of information on vote share is dampened by the weight voters put on the incumbent's type in their electoral decision (here, $\psi(p + (1-p)\beta) < 1/2$).

true effects depending on the geographical spread of observations (level of randomization, sampling frames) and the information provided (about performance or representatives' behaviors). Consequently, researchers should not automatically attribute variability in results to statistical noise or to frictions (voters' irrationality, preference bias, or collective action problem). Distinct outcomes may well be due to estimates referring to different true effects.

This variability in true effects is not due to the treatment itself. Indeed, we establish that (asymptotically) the incumbent's vote share in the treated group does not depend on the geographical spread of observations. The issue we uncover is due to variations in the average vote share in the control group. Even in similar contexts, this quantity is in fact not constant across or within sampling frames. Random assignment of treatment is not enough to keep everything constant.

We offer a recommendation to improve comparability and the accumulation of knowledge. We suggest that researchers collect data on governance outcome and the information to be disseminated for all units in their sample. Using a fully interacted model, treated units are then compared with their exact equivalent in the control group. Assuming similar context, our prescribed specification guarantees that the control group is identical within and across all sampling frames and estimates become fully comparable.

References

Acemoglu, Daron. 2010. "Theory, general equilibrium, and political economy in development economics."Journal of Economic Perspectives 24(3): 17-32.

Adida, Claire, Jessica Gottlieb, Eric Kramon, and Gwyneth McClendon. Forthcoming. "Under What Conditions Does Performance Information Influence Voting Behavior? Lessons from Benin." in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability.* Cambridge University Press.

Arias, Eric, Horacio A. Larreguy, John Marshall and Pablo Querubin. Forthcoming. "When Does Information Increase Electoral Accountability? Lessons from a Field Experiment In Mexico." in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability.* Cambridge University Press.

Ashworth, Scott and Kenneth W. Shotts. 2010. "Does informative media commentary reduce politicians' incentives to pander?." Journal of Public Economics 94(11-12): 838-847.

Ashworth, Scott. 2012. "Electoral accountability: recent theoretical and empirical work." Annual Review of Political Science 15: 183-201.

Ashworth, Scott and Ethan Bueno de Mesquita. 2014. "Is voter competence good for voters?: Information, rationality, and democratic performance." American Political Science Review. 108(3): 565-587.

Banerjee, Abhijit V., Selvan Kumar, Rohini Pande and Felix Su. 2011. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India." Working paper.

Bhandari, Abhit, Horacio Larreguy, and John Marshall. 2018. "An Empirical Anatomy of Political Accountability: Experimental Evidence from a Pre-Election Information Dissemination Campaign in Senegal." Working paper.

Bjorkman Nyqvist, Martina, Damien de Walque and Jakob Svensson. 2017."Experimental Evidence on the Long-Run Impact of Community-Based Monitoring." American Economic Journal: Applied Economics 9(1): 3369.

Boas, Taylor C., F. Daniel Hidalgo and Marcus A. Melo. Forthcoming. "Horizontal but Not Vertical: Accountability Institutions and Electoral Sanctioning in Northeast Brazil." in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability.* Cambridge University Press.

Buntaine, Mark, Sarah Bush, Ryan Jablonski, Dian Nielson and Paula Pickering. Forthcoming. "Budgets, SMS Texts, and Votes in Uganda." in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability.* Cambridge University Press.

Chong, Alberto, Ana De La O, Dean Karlan and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Party Identification."Journal of Politics 77(1): 5571.

Davis, Jonathan, Jonathan Guryan, Kelly Hallberg and Jens Ludwig. 2017. "The Economics of Scale-Up." No. w23925. National Bureau of Economic Research.

Deaton, Angus, and Nancy Cartwright. 2017."Understanding and misunderstanding randomized controlled trials."Social Science & Medicine.

de Figueiredo, Miguel F.P., F. Daniel Hidalgo and Yuri Kasahara. 2013. "When Do Voters Punish Corrupt Politicians? Experimental Evidence from Brazil."Working paper.

Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh. Forthcoming. "Do Informational Campaigns Promote Electoral Accountability?", in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability*. Cambridge University Press.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazils Publicly Released Audits on Electoral Outcomes." Quarterly Journal of Economics 123(2):703 745.

Fox, Justin. 2007. "Government transparency and policymaking." Public choice 131(1-2): 23-44.

Fox, Justin and Richard Van Weelden. 2012. "Costly transparency." Journal of Public Economics 96(1-2): 142-150.

Gottlieb, Jessica. 2016. "Greater expectations: A field experiment to improve accountability in mali." American Journal of Political Science 60(1): 143-157.

Grossman, Guy, Macartan Humphreys, and Gabriella Sacramone-Lutz. 2014. "I wld like u WMP to extend electricity 2 our village: On Information Technology and Interest Articulation."American Political Science Review 108(3): 688-705.

Grossman, Guy, Macartan Humphreys, and Gabriella Sacramone-Lutz. 2018. "Information technology and political engagement: Mixed evidence from Uganda.". Working Paper.

Grossman, Guy, Kristin Michelitch, and Carlo Prato. 2018. " Candidate entry and vote choice in the wake of incumbent performance transparency initiatives." Working Paper.

Humphreys, Macartan and Jeremy Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda Preliminary Analysis." Working paper.

Larreguy, Horacio and Shelley Liu. 2018. "The Effect of Education on Political Participation: Evidence from a Competitive Consolidating Democracy.". Working Paper.

Lieberman, Evan S., Daniel N. Posner and Lily L. Tsai. 2014. "Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya" World Development 60:6983.

Lierl, Malte and Marcus Holmlund. Forthcoming. "Performance-Based Voting in Local Elections: Experimental Evidence from Burkina Faso." in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability*. Cambridge University Press.

Olken, Benjamin A. 2007. "Monitoring corruption: evidence from a field experiment in Indonesia." Journal of political Economy. 15(2): 200-249.

Platas, Melina and Pia Raffler. Forthcoming. "Meet the Candidates: Field Experimental Evidence on Learning from Politician Debates in Uganda." in Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, and Craig McIntosh (Eds.). *Metaketa I: The Limits of Electoral Accountability*. Cambridge University Press.

Prat, Andrea. 2005. "The wrong kind of transparency." American Economic Review 95(3): 862-877.

Wolton, Stephane. 2018."Are Biased Media Bad for Democracy?." Working Paper.

A Proofs for Section 3

Proof of Lemma 1

Denote $U^e(\tau)$ the second-period expected payoff when a type τ politician is in office, $\tau \in \{b, g\}$. Using the reasoning in the main text, simple calculation yields: $U^e(b) = R$ and $U^e(g) = R + \alpha(p + (1-p)\beta) - 1$ since a bad type exerts no effort and good type exerts full effort.

From the reasoning and computation in the main text, we can rewrite the maximization problem of a type τ^I incumbent in an environment $\theta_s \in \{\beta, 1\}$ as

$$\max_{\mathbf{e^t_v}} \frac{U^e(\tau^I)}{2} + \sum_{v \in V} \left\{ \begin{array}{l} \theta_s e^1_v \Big(\alpha(\tau^I) + \frac{\phi}{V}(p + (1-p)\beta)\big(\lambda\mu_v(1, \iota^a_v) + (1-\lambda)\mu_v(\emptyset, \iota^a_v) - \pi^C\big)U^e(\tau^I)\Big) \\ +(1 - \theta_s e^1_v)\Big(\lambda\frac{\phi}{V}(p + (1-p)\beta)\big(\lambda\mu_v(0, \iota^a_v) + (1-\lambda)\mu_v(\emptyset, \iota^a_v) - \pi^C\big)U^e(\tau^I)\Big) \end{array} - \frac{(e^t_v)^2}{2} \right\} \tag{A.1}$$

Hence, the incumbent chooses the optimal level of effort for each village in isolation. □

Proof of Lemma 2

Follows directly from the reasoning in the text. □

Proof of Lemma 3

Using Equation A.1, it can be checked that a bad type's effort as a function of the environment when unaware of any treatment ($\iota^a_v = \emptyset$) is.

$$e^1_v(b, \theta_s, \emptyset) = \theta_s \times \lambda\frac{\phi}{V}(p + (1-p)\beta)(\mu_v(1, \emptyset) - \mu_v(0, \emptyset))R \tag{A.2}$$

Assuming existence, notice immediately from Equation A.2 that $e^1_v(b, 1, \emptyset) > e^1_v(b, \beta, \emptyset)$.

We first prove existence and uniqueness of the equilibrium. To do so, we denote $\overline{\mu}(e) = \frac{\pi^I(p+(1-p)\beta)}{\pi^I(p+(1-p)\beta)+(1-\pi^I)(p+(1-p)\beta^2)e}$ and $\underline{\mu}(e) = \frac{\pi^I(1-p)(1-\beta)}{\pi^I(1-p)(1-\beta)+(1-\pi^I)(1-(p+(1-p)\beta^2)e)}$. For any $e \in [0, 1]$, it can be checked that $\overline{\mu}'(e) < 0$ and $\underline{\mu}'(e) > 0$. By Bayes' rule, the incumbent's perceived voters' posterior satisfies $\mu_v(1, \emptyset) = \overline{\mu}(e^1_v(b, 1, \emptyset))$ and $\mu_v(0, \emptyset) = \underline{\mu}(e^1_v(b, 1, \emptyset))$.[12] Let $\gamma(\lambda) := \lambda\frac{\phi}{V}(p + (1-p)\beta)R$, and

[12] Observe that voters' posteriors satisfy $\mu_v(1, \emptyset) = \overline{\mu}(e^1_v(b, 1, \emptyset))$ and $\mu_v(0, \emptyset) = \underline{\mu}(e^1_v(b, 1, \emptyset))$ for ex-post performance treatment ($\iota_v = 1$). For ex-post action treatment ($\iota = 3$), treated voters' posteriors are 1 if $e^1_v = 1$ and 0

consider the function:

$$f(e;\lambda) = e - \gamma(\lambda)(\overline{\mu}(e) - \underline{\mu}(e)) \tag{A.3}$$

To prove existence and uniqueness, we just need to show that there exists a unique solution to the equation $f(e) = 0$. By the properties of $\overline{\mu}(\cdot)$ and $\underline{\mu}(\cdot)$ observe that $f_e(e;\lambda) > 0$. Further, $f(0;\lambda) < 0$ and $f(1;\lambda) > 0$. Hence, by the theorem of intermediate values, an equilibrium effort exists, is unique, and satisfies all the properties described in the text of the lemma. □

For ease of exposition, we denote $e^*(\lambda)$ the unique level of effort which solves $f(e;\lambda) = 0$. That is, $e(b,1,\emptyset) = e^*(\lambda)$ and $e(b,\beta,\emptyset) = \beta e^*(\lambda)$. The next Lemma characterizes properties of $e^*(\lambda)$.

Lemma A.1. *The equilibrium effort is strictly increasing and concave with the share of informed voters: $\partial e^*(\lambda)/\partial\lambda > 0$ and $\partial^2 e^*(\lambda)/\partial\lambda^2 < 0$.*

Proof. $f(\cdot)$ defined in Equation A.3 has the following properties: $f_\lambda(e;\lambda) < 0$ and $f_e(e;\lambda) > 0$ for all $\lambda \in [0,1]$. Thus, $\partial e^*(\lambda)/\partial\lambda > 0$ by the implicit function theorem.

Further, we can write the derivative as:

$$\frac{\partial e^*(\lambda)}{\partial\lambda} = \frac{\frac{\phi}{V}(p+(1-p)\beta)R(\overline{\mu}(e^*(\lambda)) - \underline{\mu}(e^*(\lambda)))}{1 - \lambda\frac{\phi}{V}(p+(1-p)\beta)R(\overline{\mu}'(e^*(\lambda)) - \underline{\mu}'(e^*(\lambda)))}$$

Notice thus that $\frac{\partial e^*(\lambda)}{\partial\lambda} < 1$.

Applying the implicit function theorem a second time to Equation A.3, we obtain:

$$\begin{aligned}\frac{\partial^2 e^*(\lambda)}{\partial\lambda^2}\big(1 - \lambda(\overline{\mu}'(e^*(\lambda)) - \underline{\mu}'(e^*(\lambda)))\big) =& 2\frac{\partial e^*(\lambda)}{\partial\lambda}\big(\overline{\mu}'(e^*(\lambda)) - \underline{\mu}'(e^*(\lambda)))\big) \\ &+ \left(\frac{\partial e^*(\lambda)}{\partial\lambda}\right)^2\big(\overline{\mu}''(e^*(\lambda)) - \underline{\mu}''(e^*(\lambda)))\big)\end{aligned}$$

Using the definition of $\overline{\mu}(\cdot)$ and $\underline{\mu}(\cdot)$, $\overline{\mu}''(e^*(\lambda)) = -\frac{(p+(1-p)\beta^2)(1-\pi^I)}{(p+(1-p)\beta)\pi^I + (p+(1-p)\beta^2)e^*(\lambda)(1-\pi^I)}\overline{\mu}'(e^*(\lambda))$ and $\underline{\mu}''(e^*(\lambda)) = -\frac{(p+(1-p)\beta^2)(1-\pi^I)}{(1-p)(1-\beta)\pi^I + \left(1-(p+(1-p)\beta^2)e^*(\lambda)\right)(1-\pi^I)}\overline{\mu}'(e^*(\lambda))$ □

---

if $e_v^1 < 1$ (see below). Hence, there is a mismatch between perceived and actual beliefs. Again, this is due to the difficulty associated with making the incumbent's unaware of the treatment and it could be solved by assuming that the office-holder expects an intervention to occur with probability ε, with $\varepsilon \to 0$.

Using Lemma A.1, it is immediate from the definition of $\overline{\mu}(\cdot)$ and $\underline{\mu}(\cdot)$ that $\overline{\mu}(e^*(\lambda))$ ($\underline{\mu}(e^*(\lambda))$) is strictly decreasing (increasing) with λ.

Proof of Lemma 4

As above, a bad type's effort as a function of the environment for ex-ante performance treatment ($\iota_v^a = 2$) is.

$$e_v^1(b, \theta_s, 2) = \theta_s \times \frac{\phi}{V}(p + (1-p)\beta)(\mu_v(1,2) - \mu_v(0,2))R \tag{A.4}$$

Equilibrium effort corresponds to the unique solution to $f(e; 1) = 0$. Noting that $e(b, 1, 2) = e^*(1)$, we obtain $e(b, \theta_s, 2) > e(b, \theta_s, \emptyset)$. Under the assumptions, $f(1; 1) > 0$ so $e(b, 1, 2) < 1$. Finally, $e_v^1(b, \beta, 2) - e_v^1(b, \beta, \emptyset) = \beta(e_v^1(b, 1, 2) - e_v^1(b, 1, \emptyset))$, which proves the last point of the lemma given $\beta < 1$. □

Proof of Lemma 5

Our requirement that out-of-equilibrium satisfies the D1 criterion guarantees that in any equilibrium, a good incumbent exerts maximum effort.

To see this, suppose first there is a separating equilibrium in which a good incumbent does not exert maximum effort (it can easily be checked that in any separating equilibrium, a bad incumbent exerts strictly less effort than a good incumbent). Then only a good incumbent is willing to exert maximum effort (even if voters believe only such type deviate to maximum effort) so a good type cannot exert interior effort in a separating equilibrium.

Suppose now that there is a pooling equilibrium with level of effort $e^p \in [0, 1)$. Slightly abusing notation, denote $\mathbb{P}_v(\mu_v)$ the participation of village v to the incumbent's re-election probability as a function of voters' belief: $\mathbb{P}(\mu_v) = \frac{\phi}{V}(p + (1-p)\beta)(\mu_v - \pi^C)$ using Equation 5. Denote $\hat{\mu}_v$ voters' out-of-equilibrium posteriors following a deviation to maximum effort. A good and bad types prefer full effort to e^p if and only if, respectively

$$\alpha\theta_s(1-e^p) - \left(\frac{1}{2} - \frac{(e^p)^2}{2}\right) + \big(\mathbb{P}(\hat{\mu}_v) - \mathbb{P}(\pi^I)\big)U^e(g) \geq 0$$
$$-\left(\frac{1}{2} - \frac{(e^p)^2}{2}\right) + \big(\mathbb{P}(\hat{\mu}_v) - \mathbb{P}(\pi^I)\big)U^e(b) \geq 0$$

The second inequality is never satisfied for $\hat{\mu}_v \leq \pi^I$. Hence, the set of out-of-equilibrium belief such that a bad type is willing to deviate to full effort is a strict subset of the set of out-of-equilibrium such that a good type is willing to deviate to full effort. Hence, by the D1 criterion, e^p cannot be an equilibrium level of effort.

A similar reasoning excludes semi-separating equilibrium in which one type mixes between two or more levels of effort, none corresponding to full effort.

Given a good incumbent's effort strategy, voters' posteriors satisfy $\mu_v(\omega_v^1, 4) = 0$ for all $e_v^1 < 1$ and all $\omega_v^1 \in \{0, 1\}$. This directly implies that a bad incumbent either exerts no effort or full effort (any interior level of effort comes at a cost and no electoral benefit compared to null effort).

A separating equilibrium in which a bad incumbent always exerts no effort exists if and only if

$$-\frac{1}{2} + \mathbb{P}(1)R \leq \mathbb{P}(0)R$$

The left-hand side corresponds to a bad type's payoff if he exerts full effort. He pays a cost 1/2. His re-election probability, however, increases by $\mathbb{P}(1) - \mathbb{P}(0)$ since voters believe only a good type exerts full effort. Rearranging, we obtain that a separating equilibrium exists if and only if

$$R \leq \frac{1}{2\frac{\phi}{V}(p + (1-p)\beta)} := \underline{R} \tag{A.5}$$

In turn, a pooling equilibrium exists if and only if a bad incumbent prefers to exert full effort (even though voters anticipate types play a pooling strategy and have posterior π^I after observing effort $e_v^1 = 1$) to no effort. A bad type's incentive compatibility constraint is then:

$$-\frac{1}{2} + \mathbb{P}(\pi^I)R \geq \mathbb{P}(0)R$$

Rearranging, we obtain that a pooling equilibrium exists if and only if

$$R \geq \frac{1}{2\frac{\phi}{V}(p+(1-p)\beta)\pi^I} := \overline{R} \tag{A.6}$$

For all $R \in (\underline{R}, \overline{R})$, a bad type incumbent mixes between no and full effort. Denote η the probability, he exerts full effort. Let $\mu_v^{ss}(\eta) = \frac{\pi^I}{\pi^I+(1-\pi^I)\eta}$ voters' posterior that the incumbent is a good type after observing full effort. η is uniquely defined by the following equation:

$$-\frac{1}{2} + \mathbb{P}(\mu_v^{ss}(\eta))R = \mathbb{P}(0)R$$

Since $\mu_v^{ss}(\eta)$ is strictly decreasing with η, we obtain that R is strictly increasing with R as claimed.

□

B Proofs for Section 5

For ease of exposition, we assume that for ex-ante treatments ($\iota \in \{2,4\}$), voters in the control group, like the incumbent, are aware of the treatment. This is without loss of generality and simply ensures that we can use a single formula for the true treatment effect(s) in the context of within village analysis. This assumption implies that for ex-ante performance treatment, the equilibrium level of effort is $e^*(\xi+(1-\xi)\lambda)$ with $e^*(\cdot)$ defined in the proof of Lemma 3. For ex-ante action treatment, our assumption on the office rents R imply that a bad incumbent does not exert full effort. In turn, because $(1-\xi)$ voters do not learn his level of effort, his equilibrium effort is $e^*((1-\xi)\lambda)$. For ex-post treatment, the office-holder is not aware of the treatment, and if bad, exerts effort $e^*(\lambda)$ as discussed extensively.

Proof of Proposition 1

(i) Using Equation 3 and the fact that there is a mass of voters, $E\big(S_\xi(\iota_i)|\mathcal{T}, n^+\big) = \frac{1}{2} + \psi\big((p+(1-p)\beta)(\mu_v(1,\iota_i) - \pi^C)\big)$ and $E\big(S_\xi(\iota_i)|\mathcal{T}, n^-\big) = \frac{1}{2} + \psi\big((p+(1-p)\beta)(\mu_v(0,\iota_i) - \pi^C)\big)$ for $\iota_i \in \{1,2\}$. In the first expectation, treated voters learn that the project was successful (good news) and so form posterior $\mu_v(1,\iota_i)$. In the second expectation, treated voters learn that the project failed (bad

news) and so form posterior $\mu_v(0, \iota_i)$. Notice that we do not incorporate the common valence shock (δ) in the formula since we are taking expectations and the mean of δ is 0. This is without loss of generality since δ would cancel out when subtracting the vote share in the control group.

In turn, $E\big(S_{1-\xi}(\emptyset)|\mathcal{C}\big) = \frac{1}{2} + \psi\big((p + (1-p)\beta)(\lambda\mu_v(1, \iota_i) + (1-\lambda)\pi^I - \pi^C)\big)$ if $n = n^+$ and $E\big(S_{1-\xi}(\emptyset)|\mathcal{C}\big) = \frac{1}{2} + \psi\big((p + (1-p)\beta)(\lambda\mu_v(0, \iota_v) + (1-\lambda)\pi^I - \pi^C)\big)$ if $n = n^-$. This is because, as explained in the main text, control voters are drawn from the *same* village so when informed, control voters observe the same $\omega_v^1 \in \{0, 1\}$ as the treated.

Using Equation 9, we obtain

$$VTE(\iota_v, n; \omega_v^1, \tau^I, \theta_s) = \begin{cases} \psi(p + (1-p)\beta)(1-\lambda)(\mu_v(1, \iota_i) - \pi^I) & \text{if } n = n^+ \\ \psi(p + (1-p)\beta)(1-\lambda)(\mu_v(0, \iota_i) - \pi^I) & \text{if } n = n^- \end{cases} \tag{B.1}$$

Hence, $VTE(\cdot)$ only depends on the nature of information ι_v (which translates into ξ percent of villagers treated with ι_i) and the type of information n. There is thus a unique value for each ex-post and ex-ante performance treatment ($\iota \in \{1, 2\}$) and for each good news and bad news.

(ii) Using a similar reasoning as above, we can write $E_\omega\big(S_v(\iota_v)|\mathcal{T}, n^+\big) = \frac{1}{2} + \psi\big((p + (1-p)\beta)(\mu_v(1, \iota_v) - \pi^C)\big)$ and $E_\omega\big(S_v(\iota_v)|\mathcal{T}, n^-\big) = \frac{1}{2} + \psi\big((p + (1-p)\beta)(\mu_v(0, \iota_v) - \pi^C)\big)$ for $\iota_v \in \{1, 2\}$. Turning to the control group, observe that in every village in which the project outcome is successful, the vote share of the incumbent is in expectation:

$$S_v^1 := \frac{1}{2} + \psi\big((p + (1-p)\beta)(\lambda\mu_v(1, \emptyset) + (1-\lambda)\pi^I - \pi^C)\big). \tag{B.2}$$

In turn, in every village in which the project outcome is not successful, the vote share of the incumbent is in expectation:

$$S_v^0 := \frac{1}{2} + \psi\big((p + (1-p)\beta)(\lambda\mu_v(0, \emptyset) + (1-\lambda)\pi^I - \pi^C)\big). \tag{B.3}$$

Hence, the expected vote share of the incumbent in control villages is:

$$E_\omega(S_v(\emptyset)|\mathcal{C}) = Pr(\omega_v^1 = 1|\mathcal{C})S_v^1 + Pr(\omega_v^1 = 0|\mathcal{C})S_v^0$$

Given that all villages in the control group are drawn from the same district, conditioning on $\mathcal{C}$ implies that we condition on a specific pair of type and environment as explained in the text. Hence,

$$E_\omega(S_v(\emptyset)|\mathcal{C}) = \frac{1}{2}+\psi\Big((p+(1-p)\beta)\big(\lambda Pr(\omega_v^1 = 1|\tau^I, \theta_s)\mu_v(1,\emptyset)+\lambda Pr(\omega_v^1 = 0|\tau^I, \theta_s)\mu_v(0,\emptyset)+(1-\lambda)\pi^I-\pi^C\big)\Big),$$

with

$$Pr(\omega_v^1 = 1|\tau^I = g, \theta_s) = \begin{cases} 1 & \text{if } \tau^I = g, \theta_s = 1 \\ \beta & \text{if } \tau^I = g, \theta_s = \beta \\ e(b,1,\emptyset) & \text{if } \tau^I = b, \theta_s = 1 \\ \beta e(b,\beta,\emptyset) & \text{if } \tau^I = b, \theta_s = \beta \end{cases}$$

Using Equation 6, we then obtain:

$$DTE(\iota_v, n; \tau^I, \theta_s) = \begin{cases} \psi(p+(1-p)\beta)\begin{pmatrix} \mu_v(1,\iota_v) - \lambda Pr(\omega_v^1 = 1|\tau^I,\theta_s)\mu_v(1,\emptyset) \\ -\lambda Pr(\omega_v^1 = 0|\tau^I,\theta_s)\mu_v(0,\emptyset) - (1-\lambda)\pi^I) \end{pmatrix} & \text{if } n = n^+ \\ \psi(p+(1-p)\beta)\begin{pmatrix} \mu_v(0,\iota_v) - \lambda Pr(\omega_v^1 = 1|\tau^I,\theta_s)\mu_v(1,\emptyset) \\ -\lambda Pr(\omega_v^1 = 0|\tau^I,\theta_s)\mu_v(0,\emptyset) - (1-\lambda)\pi^I) \end{pmatrix} & \text{if } n = n^- \end{cases} \tag{B.4}$$

Thus $DTE(\iota_v, n; \cdot)$ takes almost everywhere four different values, one for each pair of τ^I, θ_s. The 'almost everywhere' stands for the special case for which $e(b,1,\emptyset) = \beta$ which holds for a set of parameter values of measure zero.

(iii) The reasoning follows similar steps as point (ii) except that we now take expectations over incumbent's types since villages are drawn from multiple districts. First, note that the vote share for the treatment group remains the same: $E_{\omega,\tau^I}\big(S_v(\iota_v)|\mathcal{T}, n^+\big) = \frac{1}{2} + \psi\big((p+(1-p)\beta)(\mu_v(1,\iota_v) - \pi^C)\big)$ and $E_{\omega,\tau^I}\big(S_v(\iota_v)|\mathcal{T}, n^-\big) = \frac{1}{2} + \psi\big((p+(1-p)\beta)(\mu_v(0,\iota_v) - \pi^C)\big)$ for $\iota_v \in \{1,2\}$ since it does not depend on the state. In turn, given that all villages in the control group are drawn from the same state, conditioning on $\mathcal{C}$ implies that we condition on a specific environment. The vote share in the

control group is then

$$
\begin{aligned}
E_{\omega,\tau^I}(S_v(\emptyset)|\mathcal{C}) =& Pr(\tau^I = g)\Big(Pr(\omega_v^1 = 1|\mathcal{C}, \tau^I = g)S_v^1 + Pr(\omega_v^1 = 0|\mathcal{C}, \tau^I = g)S_v^0\Big) \\
&+ Pr(\tau^I = b)\Big(Pr(\omega_v^1 = 1|\mathcal{C}, \tau^I = b)S_v^1 + Pr(\omega_v^1 = 0|\mathcal{C}, \tau^I = b)S_v^0\Big).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
E_{\omega,\tau^I}(S_v(\emptyset)|\mathcal{C}) =& \frac{1}{2} + \psi(p + (1-p)\beta)\Big(\lambda\big(\pi^I Pr(\omega_v^1 = 1|\tau^I = g, \theta_s) + (1-\pi^I)Pr(\omega_v^1 = 1|\tau^I = b, \theta_s)\big)\mu_v(1, \emptyset) \\
&+ \lambda\big(\pi^I Pr(\omega_v^1 = 0|\tau^I = g, \theta_s) + (1-\pi^I)Pr(\omega_v^1 = 0|\tau^I = b, \theta_s)\big)\mu_v(0, \emptyset) + (1-\lambda)\pi^I - \pi^C\Big),
\end{aligned}
$$

with

$$
Pr(\omega_v^1 = 1|\tau^I = g, \theta_s) = \begin{cases} 1 & \text{if } \theta_s = 1 \\ \beta & \text{if } \theta_s = \beta \end{cases} \quad \text{and} \quad Pr(\omega_v^1 = 1|\tau^I = b, \theta_s) = \begin{cases} e(b, 1, \emptyset) & \text{if } \theta_s = 1 \\ \beta e(b, \beta, \emptyset) & \text{if } \theta_s = \beta \end{cases}
$$

Using Equation 7, , we then obtain:

$$
STE(\iota_v, n; \tau^I, \theta_s) = \begin{cases}
\psi(p + (1-p)\beta)\begin{pmatrix} \mu_v(1, \iota_v) - (1-\lambda)\pi^I \\ -\lambda\big(\pi^I\theta_s + (1-\pi^I)\theta_s e(b, \theta_s, \emptyset)\big)\mu_v(1, \emptyset) \\ -\lambda\big(\pi^I(1-\theta_s) + (1-\pi^I)(1-\theta_s e(b, \theta_s, \emptyset))\big)\mu_v(0, \emptyset)) \end{pmatrix} & \text{if } n = n^+ \\
\psi(p + (1-p)\beta)\begin{pmatrix} \mu_v(0, \iota_v) - (1-\lambda)\pi^I \\ -\lambda\big(\pi^I\theta_s + (1-\pi^I)\theta_s e(b, \theta_s, \emptyset)\big)\mu_v(1, \emptyset) \\ -\lambda\big(\pi^I(1-\theta_s) + (1-\pi^I)(1-\theta_s e(b, \theta_s, \emptyset))\big)\mu_v(0, \emptyset)) \end{pmatrix} & \text{if } n = n^-
\end{cases} \tag{B.5}
$$

Hence, $STE(\iota_v, n; \theta_s)$ takes two different values, one for each pair of θ_s.

(iv) The reasoning for the treatment group remains the same. For the control group, all villages are drawn from the different states, conditioning on $\mathcal{C}$ implies that we no longer condition on a type or

environment. The vote share in the control group is then

$$E_{\omega,\tau^I,\theta_s}(S_v(\emptyset)|\mathcal{C}) = \sum_{\tau^I\in\{b,g\}}\sum_{\theta_s\in\{\beta,1\}}\Big(Pr(\omega_v^1=1|\mathcal{C},\tau^I,\theta_s)S_v^1 + Pr(\omega_v^1=0|\mathcal{C},\tau^I,\theta_s)S_v^0\Big) \tag{B.6}$$

Using Equation B.2, Equation B.3, Equation B.6, and the martingal property of posteriors, we obtain:

$$E_{\omega,\tau^I,\theta_s}(S_v(\emptyset)|\mathcal{C}) = \frac{1}{2} + \psi\big((p+(1-p)\beta)(\pi^I-\pi^C)\big) \tag{B.7}$$

Using Equation 8, we obtain

$$ATE(\iota_v, n) = \begin{cases} \psi(p+(1-p)\beta)(\mu_v(1,\iota_v)-\pi^I) & \text{if } n=n^+ \\ \psi(p+(1-p)\beta)(\mu_v(0,\iota_v)-\pi^I) & \text{if } n=n^- \end{cases} \tag{B.8}$$

Thus, $ATE(\iota_v, n)$ then takes a unique value. □

Proof of Corollary 1

Comparing Equation B.1 and Equation B.8 yields $VTE(1, n; \cdot) = (1-\lambda)ATE(1; n)$ as claimed since a bad incumbent's effort is unaffected by the treatment.

For ex-ante performance treatment ($\iota = 2$), in the case of within village analysis, a bad incumbent's equilibrium effort in the treated village when the environment is favorable is $e^*(\xi+(1-\xi)\lambda)$ (and $\beta\times e^*(\xi+(1-\xi)\lambda)$ when the environment is unfavorable). In a country-wide analysis, a bad incumbent's equilibrium effort in treated village when the environment is favorable is $e^*(1) > e^*(\xi+(1-\xi)\lambda)$ (Lemma A.1). This then implies that $\mu_v(1, \iota_i = 2) = \overline{\mu}(e^*(\xi+(1-\xi)\lambda)) > \overline{\mu}(e^*(1)) = \mu_v(1, \iota_v = 2)$, where ι_i stands for village-level analysis, and inversely for the posteriors following project failure. Hence, $VTE(2, n; \cdot) \neq (1-\lambda)ATE(2, n)$.

Combining the analysis for ex-post and ex-ante performance treatments proves the claim. □

Proof of Corollary 2

The claim follows from the fact that the higher the proportion of successful projects in the control group, the lower the true effect (see Equation B.4 and Equation B.5). Using Table 2 in the main text, it is clear that the greatest (lowest) proportion of success occurs for the pair $\tau^I = g$, $\theta_s = 1$ ($\tau^I = b$,

$\theta_s = \beta$) so $DTE(\iota, n; g, 1) < DTE(\iota, n; b, \beta)$ for $n \in \{n^+, n^-\}$. Using the same table, it is clear that the proportion of successful projects is higher in a favorable relative to an unfavorable environment and falls between the two extremes since $\pi^I \in (0, 1)$. Hence, $DTE(\iota, n; g, 1) < STE(\iota, n; 1) < STE(\iota, n; \beta) < DTE(\iota, n; b, \beta)$ for $n \in \{n^+, n^-\}$. $\square$

Proof of Corollary 3

For ex-post performance treatment ($\iota = 1$), voters' posteriors conditional on success/failure are similar in the control and treated group. Since $\mu_v(1, 1) = \mu_v(1, \emptyset) > \pi^I > \mu_v(0, 1) = \mu_v(0, \emptyset)$, the signs directly follow from Equations B.1, B.4, B.5, and B.8. $\square$

In what follows, we always assume that $\xi < \overline{\xi} = min \in \{\frac{1}{2R(\frac{1}{2}+\frac{\phi}{V}(p+(1-p)\beta)\pi^I}, \frac{1-2R(\frac{1}{2}+\frac{\phi}{V}(p+(1-p)\beta)(\lambda+(1-\lambda)\pi^I)}{2R(\frac{1}{2}+\frac{\phi}{V}(p+(1-p)\beta)(1-\lambda)(1-\pi^I)}\}$. The assumption guarantees that a bad type under ex-ante action treatment with within-village analysis (i.e., ξ voters learn the incumbent's effort) exerts effort as if a proportion $\lambda(1-\xi)$ of voters observe project outcome (i.e. he ignores the treated voters).

Proof of Proposition 2

The claim that for all $\iota \in \{1, 2\}$ the VTE always equals zero when $\lambda = 1$ follows straightforwardly from Equation B.1. Consider now treatments administered at the village level. First, notice that when $\lambda = 1$ performance treatments have no effect on the incumbent's effort choice. As such, conditioning on good news (bad news) the incumbent's expected vote-share in the treated villages is the same as in control villages that experience success (failure). However, with the exception of within district analysis with a good incumbent under a favourable state, the control group will include both villages where the project failed and villages where it succeeded. Hence, the average vote share in the control group is different from the average vote-share in the treated villages, and the true treatment effect is not null.

Proof of Proposition 3

The proof proceeds along the same line as for the proof of Proposition 1 with two differences. First, regarding within village analysis, conditioning on good or bad news for action treatment no longer controls for project outcome in the treated village. After good news, treated voters form a posterior equal to 1 that the incumbent is a good type under our assumption. However, villagers in the control group can nonetheless observe project failures. This occurs if the environment is unfavorable. That is, after good news, the $VTE(\cdot)$ satisfies:

$$VTE(\iota_v, n^+; \omega_v^1, \tau^I = g, \theta_s) = \begin{cases} \psi(p + (1-p)\beta)(1 - (1-\lambda)\mu_v(1, \emptyset) - \lambda\pi^I) & \text{if } \omega_v^1 = 1 \\ \psi(p + (1-p)\beta)(1 - (1-\lambda)\mu_v(0, \emptyset) - \lambda\pi^I) & \text{if } \omega_v^1 = 0 \end{cases} \tag{B.9}$$

After bad news ($e_v^1 < 1$), treated voters form a posterior equal to 0 that the incumbent is a good type. However, villagers in the control group may still observe project success if the treatment is ex-post (since a bad type exerts strictly positive effort). That is, after bad news, the $VTE(\cdot)$ satisfies:

$$VTE(3, n^-; \omega_v^1, \tau^I = b, \theta_s) = \begin{cases} \psi(p + (1-p)\beta)(0 - (1-\lambda)\mu_v(1, \emptyset) - \lambda\pi^I) & \text{if } \omega_v^1 = 1 \\ \psi(p + (1-p)\beta)(0 - (1-\lambda)\mu_v(0, \emptyset) - \lambda\pi^I) & \text{if } \omega_v^1 = 0 \end{cases} \tag{B.10}$$

Second, in ex-post information treatment and ex-ante ones when $R \leq \underline{R}$, good news always indicates that the incumbent is a good type. The $DTE(\cdot)$ then becomes:

$$DTE(\iota_v, n^+; \tau^I = g, \theta_s) = \psi(p + (1-p)\beta) \begin{pmatrix} 1 - \lambda Pr(\omega_v^1 = 1 | g, \theta_s)\mu_v(1, \emptyset) \\ -\lambda Pr(\omega_v^1 = 0 | g, \theta_s)\mu_v(0, \emptyset) - (1-\lambda)\pi^I) \end{pmatrix} \tag{B.11}$$

Hence, the proportion of successful projects in control villages then only varies according to the environment. The $DTE(\cdot)$ takes two values, one for each θ_s.

For ex-ante action treatment when $R > \underline{R}$, both good and bad types produce full effort with (ex-ante) positive probability so the proportion of successful project in control villages is a function of type and environment as for performance treatment.

Bad news, in turn, always reveals that the incumbent is a bad type so using the same reasoning as above, the proportion of successful project in treated villages only vary with the environment. □

Proof of Corollary 4

In ex-post action treatment, the treated voters' posterior that the incumbent is a good type always equals 1 after good news ($e_v^1 = 1$) and 0 after bad news ($e_v^1 < 1$). Using the formulas above, the sign is always positive for good news, and negative for bad.

For ex-ante treatment, treated voters' posterior that the incumbent is good after bad news ($e_v^1 < 1$) is always zero. Hence, as above, the sign is always negative. Treated voters' posterior after good news is 1 for $R \leq \underline{R}$, π^I for $R \geq \overline{R}$, and strictly decreasing with R for $R \in (\underline{R}, \overline{R})$. To show how within district analysis can yield negative treatment effect, suppose that the incumbent is good and the environment is favorable. The $DTE(\cdot)$ then satisfies

$$DTE(4, n^+; g, 1) = \psi(p + (1-p)\beta)(\mu_v^{ss}(\eta(R)) - \lambda\mu_v(1, \emptyset) - (1-\lambda)\pi^I),$$

with $\eta(R)$ the probability that a bad type plays full effort in a semi-separating equilibrium and $\mu_v^{ss}(\eta)$ the voters' associated posterior following $e_v^1 = 1$. The $DTE(4, n^+; g, 1)$ is strictly decreasing with R and $DTE(4, n^+; g, 1) < 0$ for $R = \overline{R}$. Hence, there exists $R^+(g, 1) \in (\underline{R}, \overline{R})$ so that the sign of the treatment is negative for all $R > R^+(g, 1)$. To show how within state analysis can yield negative treatment effect, suppose that the environment is favorable. The $STE(\cdot)$ then satisfies

$$STE(4, n^+; 1) = \psi(p + (1-p)\beta) \begin{pmatrix} \mu_v^{ss}(\eta(R)) - \lambda(\pi^I + (1-\pi^I)e_v^1(b, 1, \emptyset))\mu_v(1, \emptyset) \\ -\lambda(1-\pi^I)(1 - e_v^1(b, 1, \emptyset))\mu_v(0, \emptyset) - (1-\lambda)\pi^I \end{pmatrix}$$

By the martingale property of posteriors, $Pr(\theta_s = 1)STE(4, n^+; 1) + Pr(\theta_s = \beta)STE(4, n^+; \beta) = ATE(4, n^+)$. Using Equation B.8 and the fact that the proportion of successful project in the control group is higher when the environment is favorable, we must have:

$$STE(4, n^+; 1) < \psi(p + (1-p)\beta)\left(\mu_v^{ss}(\eta(R)) - \pi^I\right)$$

Using a similar reasoning as for the $DTE(\cdot)$, there exists $R^+(g) \in (\underline{R}, \overline{R})$ so that the sign of the treatment is negative for all $R > R^+(g)$. $\square$

Proof of Corollary 5

For performance treatment, treated voters' posteriors are different in ex-ante and ex-post treatment since the incumbent's efforts vary whether he is aware of the treatment. Hence, using the definitions of the true effects from $VTE(\cdot)$ (Equation B.1) to $ATE(\cdot)$ (Equation B.8), it is clear that $\iota = 1$ (ex-post performance treatment) and $\iota = 2$ (ex-ante performance treatment) yield different effects. Further, for any performance treatment, since both types always exert strictly positive effort and voters are uncertain of the environment, voters' posteriors in both the treated and control groups are strictly interior after failed or successful project. In turn, for ex-post action treatment ($\iota = 3$), treated voters' posterior is either 1 after good news or 0 after bad news. Hence, any performance treatment yields different true effect than an ex-post action treatment.

Finally, for ex-ante action treatment, note that treated voters' posterior conditional on bad news is always zero. Hence, it is not comparable to a performance treatment. For good news, in all, but within village analysis, treated voters' posterior is $\mu_v(\omega_v^1, 4) = \frac{\pi^I}{\pi^I + (1-\pi^I)\eta(R)}$ $\omega_v^1 \in \{0, 1\}$, with $\mu_v(\omega_v^1, 4) \in [\pi^I, 1]$ as a function of R. In turn, for performance treatments, conditional on good news, $\mu_v(1, \iota) \in (\pi^I, 1)$, $\iota \in \{1, 2\}$. Hence, there exists a set of office-rents R such that $\mu_v(\omega_v^1, 4) = \mu_v(1, \iota)$, $\omega_v^1 \in \{0, 1\}, \iota \in \{1, 2\}$. This set, however, is of measure zero (and possibly a singleton). $\square$

C Proof for Section 6

Proof of Remark 1

Before proving the remark, we need to introduce some notation linking a within village analysis with its associated within district effect. Denote $\mu_\xi(\omega_v^1, \iota)$ the posterior of voters in the treated group in the treated village. Denote $\mu_{1-\xi}(\omega_v^1, \iota)$ the posterior of voters in the control group in the treated village. For ex-post treatments, $\mu_{1-\xi}(\omega_v^1, \iota) = \mu_v(\omega_v^1, \emptyset)$, with $\mu_v(\omega_v^1, \emptyset)$ voters' posterior in control *villages*. For ex-ante treatments, these voters are aware of the treatment so $\mu_{1-\xi}(\omega_v^1, \iota) \neq \mu_v(\omega_v^1, \iota)$. Denote $DTE_\xi(\iota, n; \tau^I, \theta_s)$ the true effect for within district analysis when a proportion ξ of voters is treated in a village.

We focus on performance treatment to prove the first part of the remark (a similar reasoning holds for action treatment). In this case, $\mu_\xi(\omega, \iota) = \mu_{1-\xi}(\omega, \iota)$. Using a similar reasoning as above, for performance treatment, the true effect for within village analysis equals:

$$DTE_\xi(\iota_v, n; \tau^I, \theta_s) = \begin{cases} \psi(p+(1-p)\beta) \begin{pmatrix} (\xi(1-\lambda)+\lambda)\mu_\xi(1,\iota_v) - \lambda Pr(\omega_v^1 = 1|\tau^I,\theta_s)\mu_v(1,\emptyset) \\ -\lambda Pr(\omega_v^1 = 0|\tau^I,\theta_s)\mu_v(0,\emptyset) - (1-\lambda)\xi\pi^I \end{pmatrix} & \text{if } n = n^+ \\ \psi(p+(1-p)\beta) \begin{pmatrix} (\xi(1-\lambda)+\lambda)\mu_\xi(0,\iota_v) - \lambda Pr(\omega_v^1 = 1|\tau^I,\theta_s)\mu_v(1,\emptyset) \\ -\lambda Pr(\omega_v^1 = 0|\tau^I,\theta_s)\mu_v(0,\emptyset) - (1-\lambda)\xi\pi^I \end{pmatrix} & \text{if } n = n^- \end{cases} \tag{C.1}$$

As above, the lower the proportion of successful projects, the higher the DTE.

In turn, we can rewrite the $VTE(\cdot)$ using our new notation:

$$VTE(\iota_v, n; \omega_v^1, \tau^I, \theta_s) = \begin{cases} \psi(p+(1-p)\beta)(1-\lambda)(\mu_\xi(1,\iota_i) - \pi^I) & \text{if } n = n^+ \\ \psi(p+(1-p)\beta)(1-\lambda)(\mu_\xi(0,\iota_i) - \pi^I) & \text{if } n = n^- \end{cases} \tag{C.2}$$

For $\tau^I = g, \theta_s = 1$, comparing Equation C.1 and Equation C.2, we obtain $DTE(1, n; g, 1) = \xi VTE(1, n; \omega_v^1, g, 1)$. For $\iota = 2$, the relationship does not hold and the two quantities cannot be directly compared. For every other pair, it is unclear whether $DTE(\iota, n; g, 1) \lesseqgtr VTE(\iota, n; \omega_v^1, g, 1)$. We now prove the second part of the remark for $\iota = 4$ and bad news. Recall that in this case, the incumbent exerts effort as if $(1-\xi)\lambda$ voters are informed about project outcome in the treated village. To see why the sign can change, suppose $\theta_s = \beta$ (we do not need to condition on type since only bad types produce bad news) and the project is successful in the treated village (both are positive probability events). Then, the VTE is (using the new notation and Equation B.9):

$$VTE(4, n^-; 1, \tau^I = b, \theta_s = \beta) = \psi(p+(1-p)\beta)(0 - (1-\lambda)\mu_{1-\xi}(1,\emptyset) - \lambda\pi^I) < 0$$

In turn, the associated DTE is

$$DTE_\xi(4, n^-; b, \beta) = \psi(p+(1-p)\beta) \begin{pmatrix} (1-\xi)\lambda\mu_{1-\xi}(1,4) - \lambda\beta e(b,\beta,\emptyset)\mu_v(1,\emptyset) \\ -\lambda(1-\beta e(b,\beta,\emptyset))\mu_v(0,\emptyset) - (1-\lambda)\xi\pi^I \end{pmatrix}$$

Given $\mu_{1-\xi}(1,4) > \pi^I > \mu_v(0,\emptyset)$ and $e(b,\beta,\emptyset)$ strictly increasing with β (see Equation A.4), there exists $\widehat{\beta} > 0$ such that for all $\beta < \widehat{\beta}$, $DTE_\xi(4, n^-; b, \beta) > 0$. □