Monitoring Corruption: Can Top-down Monitoring Crowd-Out Grassroots Participation?

Robert Gonzalez, Matthew Harvey, and Foteini Tzachrista*

December 23, 2020

Abstract

Empirical evidence on the effectiveness of grassroots monitoring is mixed. This paper proposes a previously unexplored mechanism that may explain this result. We argue that the presence of credible and effective top-down monitoring alternatives can undermine citizen participation in grassroots monitoring efforts. Building on Olken's (2009) road-building field experiment in Indonesia; we find a large and robust effect of the participation interventions on missing expenditures in villages without an audit in place. However, this effect vanishes as soon as an audit is simultaneously implemented in the village. We find evidence of crowding-out effects: in government audit villages, individuals are less likely to attend, talk, and actively participate in accountability meetings. They are also significantly less likely to voice general problems, corruption-related problems, and to take serious actions to address these problems. Despite policies promoting joint implementation of top-down and bottom-up interventions, this paper shows that top-down monitoring can undermine rather than complement grassroots efforts.

Keywords: Corruption, monitoring, grassroots
JEL Codes: O12, D70, D73

*Gonzalez: Department of Economics, University of South Carolina. E-mail: robert.gonzalez@moore.sc.edu. Harvey: Department of Economics, University of South Carolina. E-mail: Matthew.Harvey@grad.moore.sc.edu. Tzachrista: Department of Economics, University of South Carolina. E-mail: foteini@email.sc.edu.

1 Introduction

Corruption can flourish when citizens are disenfranchised. This has motivated policies that promote community monitoring as a means of combating corruption. In theory, bolstering grassroots monitoring may be more effective in reducing corruption than increasing top-down monitoring efforts: compared to government auditors, ordinary citizens may have better, first-hand information on the extent of corruption in their communities, as well as stronger incentives to monitor and hold officials accountable. In practice, the empirical evidence is mixed. While Bjorkman and Svensson (2009) find that government performance improves when community members engage in the accountability process, Olken (2007), Banerjee et al. (2010), and more recently Raffler et al. (2019), find limited support for the idea that increasing community monitoring results in better behavior by public officials.[1] Several mechanisms may explain the latter: grassroots monitoring has little potential in the presence of free-riding and elite capture (Olken, 2007); or if community members lack plausible deniability or means of directly reporting and punishing corrupt officials (Chassang and Miquel, 2018); or if they simply cannot properly detect corruption (Olken, 2009).[2]

This paper proposes an alternative channel: the crowding-out of individuals' incentives to participate in community monitoring resulting from effective top-down monitoring alternatives. In other words, we argue that a potential explanation for the ambiguous findings in the literature is that the presence of effective and credible top-down monitoring can undermine citizen participation in grassroots monitoring efforts.

We explore this mechanism empirically by building on seminal work by Olken (2007). Olken (2007) conducted a randomized controlled field experiment involving over 600 Indonesian road-building projects. These projects were vulnerable to corruption in the form of

[1]Other examples of papers studying a similar question and finding both positive and mixed results on the effect of community monitoring are: Pandey et al. (2009); Barr et al. (2012); Pradhan et al. (2014); Banerjee et al. (2018); Gonzalez (2020)

[2]Elite capture refers to measures taken by officials that potentially thwart the monitoring ability of citizens. For example, taking over complaints sent to an anti-corruption hotline. In the case of Olken (2007), elite capture occurred via the funneling of invitations to accountability meetings to individuals that were sympathetic to the village heads.

missing expenditures in the road-building projects. Villages were randomly assigned into three possible monitoring interventions: audits conducted by a central government agency, invitations to village accountability meetings, and a combination of invitations and anonymous comment forms. The last two were grassroots interventions while the audit treatment was a top-down, centralized approach. In this work, Olken (2007) concluded that although the audit intervention was quite successful in reducing missing expenditures, the grassroots interventions had limited success.

We take advantage of the fact that these treatments were independently assigned (and assignment was common knowledge to villagers), to explore how community monitoring behavior among villagers and missing expenditures respond to the grassroots interventions in the presence (and absence) of audits. In spite of the small sample size in the original experiment, we uncover that the effectiveness of bottom-up monitoring is significantly undermined by whether the village also had an audit intervention in place. Specifically, we find that grassroots monitoring leads to a statistically significant decrease in the share of missing expenditures of 8 to 10 percentage points in non-audit villages while the effect is close to zero in magnitude in audit villages. Interestingly, the grassroots effect in non-audit villages is comparable to the effect of the audits. In fact, when analyzing a purely grassroots or purely top-down monitoring strategy, we find that grassroots monitoring is (i) as effective as auditing in reducing corruption, and (ii) almost three times as cost-effective as auditing.

We find considerable evidence of audits undermining participation: individuals are less likely to attend, talk, and actively participate in accountability meetings if they live in an audit village. They are also significantly less likely to voice general problems, corruption-related problems, and to take serious actions to address these problems. In contrast, individuals were 5% more likely to attend accountability meetings, about 6% more likely to participate and talk during these meetings, 14% more likely to voice project-related problems, and 27% more likely to voice corruption-related problems in non-audit villages relative to audit villages. Villages are also more proactive and responsive to community monitoring in the absence of audits: the likelihood of taking serious actions in response to problems raised during the accountability meetings more than doubled in non-audit villages.

In a set of additional results, we find that the grassroots intervention seems to be less disruptive in terms of substitutions across corrupt offenses. Specifically, although audits lead officials to substitute from theft to nepotism, as documented by Olken (2007)), grassroots interventions do not. In the absence of audits, we find no evidence that family members of project heads or village officials are more likely to be employed if a grassroots intervention is in place. Differences in the nature of the monitoring technologies (external centralized audits *versus* "internal" participatory monitoring) can potentially explain this differential results on cross-corruption substitutions.

We proceed by presenting an illustrative model to help understand the mechanisms behind the documented drop in participation in the presence of an audit. We discuss the following mechanisms: (i) non-pivotal participation: the perception that an individual's marginal participation in the accountability process is non-pivotal in the presence of an effective audit, (ii) retaliation costs: audits can affect how individuals perceive the potential for retaliation from publicly voicing complaints, (iii) different public returns to audits and participation: benefits in the quality of roads from lower malfeasance can differ depending on whether this was achieved via audits or grassroots interventions,[3] and to a lesser extent, (iv) whether the distribution of pro-social norms across villages may vary with audits. Despite data limitations, we present quantitative and contextual evidence suggesting that the non-pivotal participation channel is a plausible explanation for why top-down monitoring can depress participation.

We note that this paper cannot answer whether top-down and bottom-up monitoring are substitutes or complements of each other. The data simply do not allow to fully answer this question: since the same audit likelihood and intensity was implemented regardless of

[3]For instance, audits and grassroots participation can lead to an equivalent overall drop in malfeasance. However, the drop in audit villages may result mostly from lower malfeasance in the purchase of materials whereas grassroots participation may mostly affect labor theft.

community monitoring or not, one can only answer how community monitoring responds to the presence of audits but not the opposite. In spite of sample size limitations, this paper presents evidence that top-down monitoring in the form of an external, high intensity, and credible audit, lowers the corruption-deterring effect of community monitoring and depresses participation by ordinary citizens in accountability efforts.

This paper contributes to a broad literature on the effectiveness of monitoring strategies (e.g., Olken (2007), Ferraz and Finan (2008), Bjorkman and Svensson (2009), Serra (2011), Callen and Long (2015)). Specifically, this paper sheds light on the question of whether policies that combine top-down and bottom-up monitoring are effective. We show that an unintended consequence of these policies is that top-down monitoring can actually undermine rather than complement grassroots efforts. With this in mind, this paper adds to recent work delving deeper into the effectiveness of community interventions and what mechanisms can explain the pitfalls of some of these interventions (e.g., Raffler et al. (2019)).

Our paper shows that the crowding out of citizens' incentives to engage and participate can result from mechanisms other than the typical material incentives previously proposed in the literature (Bowles and Polania-Reyes, 2012; Gneezy et al., 2011). Instead, we show that an efficient government institution can crowd out community incentives to monitor. This latter evidence adds to a smaller but important literature documenting how centralized formal institutions can undermine rule following and pro-social norms (Tabellini, 2008; Lowes et al., 2017).

The paper proceeds as follows: section 2 describes the background, experimental design, and data used. Section 3 provides results on corruption and evidence on the participation effect of the top-down intervention as well as additional results on substitution across forms of corruption and a policy-related discussion. Section 6 concludes the paper.

2 Background

This section provides a self-contained summary of the setting, data, and experimental design but refer to Olken (2007) for a more detailed description.

2.1 Setting and Experimental Design

Olken (2007) studies over 600 village infrastructure projects from the Kecamatan Development Project between 2003 and 2004 in two of Indonesia's most populous provinces: East Java and Central Java. As part of the program, each village proposes the use of funds for small-scale infrastructure and micro-finance projects. The infrastructure projects generally involve the construction of roads ranging from 0.5 - 3 kilometers both within and between villages. Upon receiving funding, the community plans construction and elects the implementation team in charge of material procurement, labor hiring, and project completion.

As part of the study, villages were randomly assigned into three possible monitoring interventions: (i) audits conducted by Badan Pengawasan Keuangan dan Pembangunan (BPKP)—a government auditing agency, and two interventions designed to increase community participation: (ii) invitations to accountability meetings, and (iii) invitations plus anonymous comment forms. Invitations allow for more representative participation while comment forms allow villagers to anonymously voice concerns with the project's handling without fear of retribution. The overall distribution of treatments was: 282 villages (about 47%) were assigned to the audit treatment. Of those, 189 (about 31%) were also assigned to either the invitations or invitation plus comment forms treatment.[4]

We note that all three treatments were assigned independently, so audit and non-audit villages are equally likely to receive the grassroots treatments. We proceed with a summary of the interventions.

[4]These numbers are presented in Table B9 in the Online Appendix.

Audits In the audit treatment, villages learned that the projects would be audited with probability one by the independent government audit agency. For reference, the audit probability in control villages was 4%. The villagers were also informed that the auditors would send the results to the central government and project officials, in addition to presenting them to the villagers in town meetings. Audits occurred in two phases: (i) about $1-4$ months after the start of construction, and (ii) about 7 months after construction was complete. After each phase, the audit results were presented to the villagers in open village meetings. Randomization for the audits treatment was done at the subdistrict level to avoid audit spillovers into nearby villages within the subdistrict.

Invitations and Invitations Plus Comment Forms In all villages, three accountability meetings took place after 40, 80, and 100 percent of project funds were spent. During these meetings village and project heads present an accountability report and explain how all funds were used. In practice, village heads invite supporters and members of the village elite to the meetings. To increase participation, the invitations treatment used schools and village heads to distribute either 300 or 500 invitations throughout the village days before each of the three meetings. This equates to about one in every two households receiving a formal, written invitation. The two forms of distribution—village head or schools—were used to test elite capture.[5] In the case of the invitations plus comment forms treatment, the invitations included an anonymous comment form asking for villagers' opinions on the project. Two days before each village meeting, enumerators collected the comment forms in a sealed box. During the meeting, the enumerator read a summary of the comment forms and a sample of the free-response questions. Randomization for both treatments was done at the village level.

Olken (2007) concludes that although the audit intervention was quite successful at reducing missing expenditures, the grassroots interventions were generally ineffective. However,

[5]Distribution via schools is less likely to suffer from elite capture as individuals potentially receive invitations from their children rather than through village heads.

he shows evidence that this was the result of elite capture via the funneling of invitations by village heads to individuals that were sympathetic to them in the community.

We explore an alternative mechanism that can potentially explain the weak effect of the grassroots interventions. We hypothesize that in villages where an audit is implemented *along with* the grassroots interventions, there will be crowding out of incentives to participate in the monitoring process (i.e., attend meetings, voice concerns, submit complaints, etc.) as there is an effective alternative. This is particularly plausible considering that villagers know whether there is also an audit and results from these audits are shared publicly as part of the accountability meetings. Conversely, in villages where only the participation interventions were implemented, we expect to find that grassroots monitoring is more successful in reducing missing expenditures. Recall that the randomization of the grassroots interventions was independent from the audit randomization. This feature allows us to experimentally explore how community monitoring incentives and missing expenditures respond to the grassroots interventions in the presence (and absence) of audits.

2.2 Data

We use a combination of the four sources of data collected by Olken (2007): (i) A village-level survey containing demographic characteristics of the villages as well as project implementation team characteristics, (ii) a household survey containing basic household information, such as the number of family members, social, religious, and government activities, (iii) a meetings survey collecting information on attendance, participation, as well as a count of the number of issues raised, whether those issues were related to corruption, and whether participants took serious actions to resolve the issues,[6] (iv) an engineering survey used to create the primary measure of corruption in the analysis.

[6] At each meeting, the enumerator kept track of attendees who actively participated in the discussions with attendees categorized as elites and non-elites. The elite attendees held official positions in the village (or the project) or were described as informal village leaders by locals. The non-elite attendees were not serving the village in any of these capacities. Serious actions include the replacement of suppliers or village officials, reimbursements for missing expenses, and further internal and external investigations into the issues.

For all results in the paper, we combine the two interventions: invitations and invitations plus comments into a single variable equal to one if a village was assigned to either of these interventions. Since both variables were randomly assigned independently of each other, the combined measure can be thought of as randomly assigned as well. We use this combined definition since the focus of the paper is in bottom-up monitoring as a whole and because it allows greater statistical power compared to separate measures of monitoring.[7] "Invitations plus comment forms" refers to the combined measure hereafter.

In order to assess the validity of the randomization after combining the two bottom-up monitoring variables, Table 1 presents summary statistics for various village characteristics. We present results separate by whether a village is in the combined invitations plus comment forms treatment or not. Specifically, column (1) presents summary statistics for control villages, column (2) for the invitation and comment forms treatment that was distributed via schools, and column (3) for the invitation and comment forms treatment that was distributed via neighborhood heads. Columns (4)-(6) present the results from mean comparison tests across the three groups. Overall, averages are similar across control and treatment villages suggesting that the sample is well balanced across treatment status. There are some exceptions, however. There are statistically significant differences in population size, village head age and salaries between treated and control villages (columns (4)-(6)). We note that although individual tests may yield statistically significant differences in some instances, tests for the joint significance of all variables in Table 1 fail to be rejected in all cases.[8] Moreover, we replicate all of our main results controlling for all variables listed in Table 1 and they remain robust and quantitatively similar to the experimental results.[9]

The measure of corruption used is the percent missing expenditures. This is defined as

[7]Results separating invitations and invitations plus comment forms are qualitatively similar and can be shown upon request.

[8]P-values reported in the last row of Table 1. These p-values come from the joint significance test of all coefficients obtained from a Probit model using all variables in Table 1 as controls to predict school distribution restricting sample to school distribution and control villages (column (4)), predict neighborhood head distribution restricting sample to neighborhood head distribution and control villages (column (5)), and predict school distribution restricting sample to school or head distribution (column (6)).

[9]Refer to Tables B10, B11, and B12 in the Online Appendix.

the difference between the log of reported expenses minus the log of expenses calculated in the engineering survey. This survey allowed obtaining an independent measure of actual expenses by estimating quantities of materials used, a worker survey to measure wages paid, and a supplier survey to measure prices of the materials (Olken, 2007).

Table 2 presents summary statistics on this measure separately by treatment type. As in Table 1, we further split the invitations and comments treatment by distribution method (school vs. neighborhood head). A simple analysis of the summary statistics yields a set of interesting findings that serve as preamble for the main results of our paper. On average, purely control villages reported missing expenses of about 30.3 percent. In purely audit villages (i.e., audit present but not grassroots intervention), percent missing expenses dropped down to about 19.2 percent, suggesting a 11 percentage point drop in missing expenses relative to purely control villages. Since the focus of our paper is on how the effect of the grassroots interventions varied with the presence of an audit, columns (4) and (5) present mean comparison tests between missing expenses in grassroots villages and control villages separately by audit status. First, notice that percent missing expenses were largely unaffected in villages with distribution via neighborhood heads (column (5)). As Olken (2007) explains, this is likely indication of elite capture rendering the intervention ineffective. School distribution, on the other hand, minimizes elite capture and thus leads to a more effective response as shown in Olken (2007). However, column (4) shows a very interesting pattern once we delve deeper into the effect of the school-distributed grassroots intervention by audit status. The grassroots intervention when distributed via schools is quite effective but *only* when it is implemented by itself (i.e., in the absence of an audit). Note in the first row of column (4) that the intervention leads to a statistically significant drop in missing expenses of about 9 percentage points (0.212 vs 0.303). This is in fact comparable to the drop obtained from the purely audit strategy (0.192 vs 0.303 in column (1)). However, when an audit is in place, the effectiveness of the grassroots intervention is quite limited: the difference in missing expenses between control and school-distributed treatment villages is about 0.9 percentage points and statistically insignificant (column (4)). This is also clear by simply looking at the averages, notice that average missing expenses in audit-only villages are around 19.2 percent. Adding the school-distributed grassroots intervention on top of the audits does not move that average in any meaningful way (20.1 vs 19.2 percent).

We note that this can only be explained by the grassroots interventions responding to the presence of an audit and not the other way around. Recall that the audits were performed with 100 percent probability; therefore, the likelihood of being audited did not change based on the presence of a participation intervention. This also applies at the intensive margin: the format of the audits used to uncover malfeasance was standard across all villages, the auditors were external players, and the presence of accountability meetings was not a factor in how the audits were implemented.[10]

3 Results

This section presents results on the effect of the grassroots interventions on missing expenditures and on various measures of participation in accountability meetings, by audit status.

3.1 Effect of Bottom-up Monitoring on Missing Expenditures by Audit Status

We start by exploring how the effect of the grassroots intervention on missing expenditures varies with audit status. Since the issue of elite capture highlighted in Olken (2007) is another potential driver of the grassroots experiment's null effect, we focus on villages where the invitations and comment forms were distributed via schools. Results from invitations distributed via schools are unlikely to suffer from elite capture as individuals received the invitations directly from their children and hence bypassing village officials. This ensures that any null effect we find from the community meetings treatment is not due to elite

[10] As noted in Olken (2007), audits consisted of "inspections of the project's financial records and a field inspection of the construction activities".

capture affecting the composition of participants.[11] Specifically, we estimate the equation:

$$Y_{ij} = \alpha_1 + \alpha_2 IC_{ij} \times School_{ij} + \alpha_3 IC_{ij} \times (1 - School_{ij}) + \delta_j + \epsilon_{ij} \quad (1)$$

separately by whether village i had a government audit or not. Y_{ij} is the outcome of interest in village i, in subdistrict j. IC_{ij} combines both treatments of bottom-up monitoring: the invitations and/or the comment forms. $School_{ij}$ denotes whether the invitations were distributed via schools. While δ_j denotes a subdistrict fixed effect. Standard errors are clustered at the village level given that the randomization for the participation interventions was done at this level. Coefficient α_2 which gives the effect of the participation interventions when distributed via schools is our coefficient of interest.[12]

For our main results, we present both equation (1) separately by audit status and the fully interacted model that includes the interaction between audit status and the participation interventions.[13] Since we are interested in estimating the effect of the grassroots interventions, we use subdistrict level fixed effects (strata level) and clustering of the standard errors at the village level (randomization level). However, note that since the treatment assignment strata differed between the audit and invitations treatments–randomization of audits was done at the subdistrict level–we are unable to estimate an audit effect in the fully interacted model since the audit treatment is constant within subdistricts. With this in mind, we still present results using different specifications of the fully interacted model which we discuss later in subsection 5.1.

Columns (1)-(3) of Table 3 presents the estimates of α_2 from equation (1) using percent missing expenditures in roads as the outcome variable. We focus our attention on road

[11] Appendix Table A2 still presents results without differentiating distribution method. The results are qualitatively similar to the main results presented in this section. However, they are less precise as expected.

[12] We estimate this form of equation (1) instead of the more conventional $Y_{ij} = \alpha_1 + \alpha_2 IC_{ij} + \alpha_3 IC_{ij} \times School_{ij} + \delta_j + \epsilon_{ij}$ in order to obtain the effect by school distribution directly from the coefficient α_2 in equation (1). Also note that one cannot estimate the interacted model: $Y_{ij} = \alpha_1 + \alpha_2 IC_{ij} + \alpha_3 School_{ij} + \alpha_3 IC_{ij} \times School_{ij} + \delta_j + \epsilon_{ij}$ since all school distribution is by default in the invitations treatment.

[13] i.e., $Y_{ij} = \alpha_1 + \alpha_2 IC_{ij} \times School_{ij} + \alpha_3 IC_{ij} \times (1 - School_{ij}) + \alpha_5 Audit_j + \alpha_6 Audit_j \times IC_{ij} \times School_{ij} + \alpha_7 Audit_j \times IC_{ij} \times (1 - School_{ij}) + \delta_j + \epsilon_{ij}$

projects as they account for more than 75% of all project expenses. However, Appendix Table A3 replicates these results using missing expenses in roads and ancillary projects as the outcome variable. Column (1) presents the effect of the participation interventions on the pooled sample. Columns (2) and (3) present results separating the analysis by the audit status of the village.

Notice in column (1) that in the pooled sample, the participation interventions have a relatively small and insignificant effect on missing expenditures. These are essentially the results shown in Olken (2007). However, once we perform the analysis separately for audited and non-audited villages, a pattern that is masked in the pooled sample emerges. Column (2) shows that in the absence of top-down monitoring, the effect of the community monitoring experiment is considerably larger than when an audit is present (Columns (3)). Specifically, grassroots interventions lead to a statistically significant 8.3 percentage point drop in missing expenses when there is no audit present. In villages where an audit is being simultaneously implemented, that effect becomes statistically zero (Column (3)).

In order to compare the difference in the grassroots effect between audit (columns (3)) and non-audit (columns (2)) villages, column (4) of Table 3 presents the fully interacted model that includes the interaction between audit status and the participation interventions.[14] We focus on the interaction term in column (4). Note that despite the small sample size and the number of parameters, we still find a statistically significant difference in the participation effects by audit status.

Overall, the results in columns (1)-(4) of Table 3 provide considerable evidence that the effect of the community interventions in villages where this was the only monitoring intervention was economically and statistically significant while this same effect was close to zero in villages where an effective and credible government audit was in place. These results hold when we use alternative measures of the outcome variable (Appendix Table A3) and alternative fixed effects and clustering (Appendix Table A1). Similarly, our results hold

[14]Note that we cannot estimate the effect of audits in the absence of the grassroots interventions (α_5) since audits are constant within subdistricts.

when using randomization inference to calculate p-values. These results are presented in Tables A1-A4 in the Online Appendix. Columns (5) and (6) use different levels of clustering and fixed effects that allow comparing a purely audit strategy *versus* a purely grassroots strategy. These results are discussed in section 5.1 below.

We proceed by exploring whether this pattern can be explained by the presence of audits having a differential effect on participation and engagement in community monitoring efforts.

3.2 Grassroots Participation by Audit Status

The previous section documents a clear drop in the corruption-deterring effect of the grassroots interventions when there is an audit in place. This section explores whether this can be explained by audits depressing participation and other measures of citizen engagement. Specifically, we estimate the following equation:

$$Y_{mij} = \alpha_1 + \alpha_2 IC_{ij} + \Omega_m + \delta_j + \epsilon_{mij} \tag{2}$$

where Y_{mij} is a measure of participation in meeting m, in village i, in subdistrict j. Ω_m is a meeting fixed effect to control for meeting-specific characteristics. The remaining terms are defined as in equation (1). As before, we estimate equation (2) separately by audit status and compare the effect of the grassroots intervention α_2 across these two conditions. We cluster standard errors at the village level given that the intervention IC_{ij} was randomized at this level.

Note that we do not estimate equation (2) separately by distribution method as we did for equation (1). The primary reason is that when one separates by treatment-distribution bins, some of the participation outcomes vary little within bins since their frequency is relatively rare. For instance, the pooled sample average share of corruption-related problems discussed in the meetings, and whether serious responses were taken in the meetings are 0.06 and 0.026, respectively, for the entire pooled sample. Therefore, we focus the discussion of our results on the impact of the invitations and comment forms intervention regardless of how they were distributed. Nonetheless, Tables A4 and A5 in the Appendix present results using distribution by school. The results are qualitatively similar to the main results presented in this section although less precise as expected.[15]

Table 4 presents estimates of α_2 in equation (2) in the pooled sample (columns (1) and (4)) and separately by audit status (columns (2),(3),(5), and (6)) for two measures of participation: the number of non-elite who attend the meetings and the number of non-elite who talk and voice concerns at the meetings. Non-elite are defined as individuals that have no official position in the village or the projects (Olken, 2007). Column (1) shows that the invitations and comments intervention significantly increased attendance by about 11.8 more individuals. The effect was about 12.9 more individuals attending in the absence of an audit (column (2)), while the presence of an audit depressed the effect on attendance by about 19% to 10.5 more individuals. Relative to average attendance, this means that the invitations and comment forms increased attendance by about 54% in non-audit villages and by about 43% in audit villages.[16]

Looking at more active measures of participation in columns (4)-(6), we find even more striking differences. In column (5), relative to the average, we document a statistically significant 39% increase in the number of non-elite who talk at the meetings (0.344 relative to average of 0.881). Once there is an audit in place, the effect of the intervention drops by almost 60% and becomes statistically insignificant (0.344 *versus* 0.143).

Given spacing constraints, we do not present the fully interacted model as in column (4) of Table 3, however, the row labeled "P-value (No audit=Audit)" provides the p-value obtained from comparing the grassroots effects across audit and non-audit villages.[17] Note that although we find an economically large difference in attendance and participation across

[15]Similarly, Tables B5 and B7 in the Online Appendix present p-values calculated using randomization inference. The results do not vary significantly from the results presented in this section using conventional inference methods.

[16]Dependent variable means presented in all tables.

[17]Specifically, this refers to the p-value on the term α_4 in the model $Y_{mij} = \alpha_1 + \alpha_2 IC_{ij} + \alpha_3 Audit_j + \alpha_4 Audit_j \times IC_{ij} + \Omega_m + \delta_j + \epsilon_{ij}$

non-audited and audited villages, the difference is not significant at conventional levels. In the case of participation at the intensive margin (columns (5) and (6)), statistical significance is near conventional thresholds and it is actually significant using alternative clustering (e.g., districts), and after adding village-level controls which typically increase precision in an experimental setting. Overall, it is important to consider that despite the lack of statistical significance in the difference, we consider the stark contrast between the magnitude and precision of the estimates in non-audited villages (column (5)) *versus* audited villages (column (6)) as valuable evidence of the detrimental effect of audits on participation.

Table 5 presents estimates of equation (2) using more direct measures of accountability. These measures address issues directly related to the management of the projects. Specifically, we examine the effect of the participation intervention on the number of problems voiced, and whether the problems were related to corruption. We also look at whether village officials are more responsive to the complaints of the community by looking at whether a serious response was taken to the problems raised in the meetings. These responses included replacing a supplier or village official, returning funds, internal village investigations, among other things (Olken, 2007). As in previous results, we split the analysis by the audit status of the village and focus on whether the participation effect is affected by the simultaneous presence of an audit.

Column (1) presents the estimated effect for the pooled sample. The invitations and comments intervention increases the number of problems raised by about 0.088 more issuer per meeting. Columns (2) and (3) show that when the effect of invitations and comment forms is estimated separately for audited and not audited villages the effects differ substantially. Invitations and comment forms increased the average number of problems discussed in a meeting by 0.188—a 16% increase relative to the baseline average—and this effect is statistically significant at a 5% significance level. Invitations and comment forms actually have a small and statistically insignificant effect on the number of problems discussed in a meeting in audit villages (column (3)).

When looking at corruption problems in particular, the pooled sample estimates show that invitations plus comment forms significantly increased any corruption related problems discussed as documented by Olken (2007). When we dissect the analysis by audit presence in columns (5) and (6) it is clear that the participation effect is entirely driven by non-audited villages. In particular, in villages without an audit, the likelihood that individuals voice issues related to corruption in the projects increases by 2.7 percentage points. This is close to a 36% increase relative to the average proportion in the sample. The findings are not as precise if there is an audit in place (column (6)).

Regarding any serious response taken during a meeting, we do not find an economically or statistically significant effect of the participation intervention in the pooled sample (column (7)). When we split the results for audited and not audited villages, we uncover that the pooled analysis is masking an interesting pattern. In non-audited villages, there is a highly significant response to the participation intervention (column (8)). The likelihood of a serious response taken increases by 2.1 percentage points and this effect is statistically significant at a 5% significance level. This represents more than a 50% increase relative to mean levels. On the other hand, invitations and comment forms have a null effect on serious responses taken in audited villages and this effect is not statistically significant.

As in Table 4, we present the p-values obtained from comparing the grassroots effects across audit and non-audit villages. In the case of problems addressed during the meetings and the likelihood of a serious response taken, we find a statistically significant difference between the effects in audited and non-audited villages using conventional levels. In the case of the likelihood of raising corruption-related issues, we do not find a statistically significant difference. However, it is important to consider that the relative findings in columns (5) and (6) still provide valuable evidence suggesting that the effect of the participation interventions seem to be more precise in the absence of an audit.

In summary, this section uncovers two key findings. First, the presence of an audit seems to inhibit the corruption-deterring effect of grassroots interventions. In villages with an audit, the invitations and comments intervention leads to a weak response in missing expenses whereas the intervention is quite effective in non-audited villages. In fact, non-audited villages with a grassroots intervention in place document a sharp drop in missing expenses. We then explore what can explain this differential effect. We uncover a similar differential effect of the invitations and comments intervention that depends on whether an audit is present in the village. Specifically, we find that in villages where an audit is in place, individuals are less likely to attend and talk during the accountability meetings. Compared to their counterparts in non-audited villages, they are also significantly less motivated to voice general problems, corruption-related problems, and to push their village to take serious actions to address these problems. To put in perspective the differential response, the effect of the participation interventions in non-audited villages is more than six times larger than in audited villages for the number of problems raised and almost twice as large for whether any serious response was taken.

The next section explores, both conceptually and empirically, how the presence of a credible, external audit can depress participation in the accountability process.

4 How can Audits lower Participation?

This section presents a framework to help understand the mechanisms behind the documented drop in participation in the presence of an audit. We describe the problem of an individual deciding whether to participate in community monitoring or not. We assume that the individual is an expected utility maximizer. In the absence of corruption the individual receives public gain $\xi > 0$, this can be thought of as the utility gain from a higher quality road or public good. Participation carries a cost $c > 0$ but can bring private utility gain $\lambda \geq 0$ which can be interpreted as "warm glow" received from engaging in pro-social behavior. In deciding whether to participate, the individual assesses that the probability that corruption will be deterred conditional on his participation $p = \{0, 1\}$ and audit intensity

$a \in [0, 1]$ is given by $\phi_{a,p}$. The individual participates if:

$$\phi_{a,1}(\xi + \lambda - c) + (1 - \phi_{a,1})(\lambda - c) \geq \phi_{a,0}\xi \tag{3}$$

Rearranging, the individual participates if $\lambda + \Delta\phi_a\xi \geq c$ where $\Delta\phi_a = (\phi_{a,1} - \phi_{a,0})$. Intuitively, the individual participates if private gains from participation and net public benefits from deterrence are higher than the cost of participation. Assuming that pro-social norms within the community are distributed with probability function $F(\lambda)$, then we can expect the probability of participating for a given individual to be $1 - F(c - \Delta\phi_a\xi)$. Note that higher costs decrease the probability of participation while a higher likelihood of deterrence and higher public benefits from no corruption make participation more likely. However, a more suitable question for our context is how does this probability of participation change given an increase in audit intensity?

To simplify the analysis and to follow our setting closely, assume that audit intensity a is either 0 (no audit) or 1 (full audit). Participation will be lower in the presence of an audit if:[18]

$$1 - F(c - \Delta\phi_1\xi) \leq 1 - F(c - \Delta\phi_0\xi) \tag{4}$$

$$\Delta\phi_1 \leq \Delta\phi_0 \tag{5}$$

where $\Delta\phi_1 = \phi_{1,1} - \phi_{1,0}$ and $\Delta\phi_0 = \phi_{0,1} - \phi_{0,0}$. In words, if the individual believes that the marginal effect of his participation on stopping corruption in the presence of an audit ($\Delta\phi_1$) is small, then he will be less likely to participate in the monitoring process. Intuitively, the individual will be less likely to participate if he perceives his participation to be trivial or redundant as the outside audit intensifies.[19]

[18] Refer to the Appendix for an extension of the analysis that allows p to be continuous and allows λ and c to depend on participation p.

[19] An alternative framing using expression (5) is that when deciding to participate, the individual will ask himself: by how much will the probability of no corruption change by my participation? With an audit it will change by $\Delta\phi_1$, without one, it will change by Δphi_0. The individual will not participate if Δphi_1 is

In a context where the outside audit is credible, effective, and its implementation is common knowledge among the community; it is plausible that individuals perceived $\Delta\phi_1$ to be zero (i.e., their "vote" or marginal participation in the presence of an audit is non-pivotal in deterring corruption). We refer to this as the "non-pivotal participation" mechanism. In such cases, expression (5) holds trivially. This can explain why we document a drop in attendance and other participation measures when a government audit is available in the community. Simply put, if people generally perceive that their marginal contribution is close to zero when there is a credible monitoring alternative; then they will be less likely to participate in collective action.

We proceed by discussing alternative explanations. Using our simple model as reference, participation can also drop in the presence of an audit if either costs c, public benefits ξ, or the distribution of pro-social norms $F(.)$ vary with audit presence. If that is the case then differences in participation across audit and non-audit locations can be explained by these differences and not the non-pivotal participation mechanism discussed above.

One can think of participation costs as the physical costs of attending the meetings plus the potential retaliation costs from attending and publicly voicing complaints. The presence of an audit will not affect physical costs but can affect perceptions about potential retaliation costs. For instance, the presence of an external third-party in the form of central government auditors may increase the perception that individuals are better shielded from retaliation. In such cases, audits will increase the likelihood of participation. We cannot observe perceived retaliation costs, however, our results showing that audits depress participation suggest that this mechanism is either unlikely or that any positive effects in participation resulting from lower retaliation costs are being overshadowed by the negative effect of the non-pivotal participation mechanism.

Public benefits ξ should not, in principle, depend on whether there is an audit or not. Intuitively, the quality of a road built without any malfeasance should be the same regardless

smaller.

of how this was achieved (audit, citizen participation, honest officials, etc.). In practice, however, malfeasance may be only partially deterred so that ξ might depend on how it was deterred. For example, the threat of an audit might lead to malfeasance in labor costs which may be harder to detect ex-post (ghost workers, overreporting wages, etc.) but not on material costs. A grassroots intervention may instead lead to malfeasance in material costs which are harder to assess by the community but not on labor costs since the laborers are recruited from within the community and have a stake in whether they are being underpaid. Therefore, overall levels of theft and malfeasance may be deterred by the same degree in both cases but in different ways. The main concern is whether this translates into different gains in road quality ξ. If ξ differs by audit status then the probability of participation will be lower in the presence of an audit if $\xi_1 < \xi_0$.[20] Intuitively, a lower ξ in audit villages makes fighting against corruption via more participation less enticing. In principle, this can explain the documented drop in participation in audit villages. However, note that if the non-pivotal participation mechanism is at play then whether ξ_1 differs from ξ_0 is irrelevant in explaining whether audits depress participation.[21]

To provide more concrete evidence on this channel, Table 6 presents results comparing several project characteristics across audit and grassroots monitoring villages conditional on the level of missing expenditures. The goal is to see whether conditional on the same level of corruption, road quality (as proxied by several characteristics) varies between audit and grassroots villages. Evidence of significant differences in characteristics/quality given the same level of corruption can indicate that public benefits ξ can depend on the monitoring technology used. In columns (1)-(6), we focus on the coefficient on the interaction term. We find no evidence that project size, the share of expenses in sand, rocks, and unskilled

[20]From expression (5), the probability of participation will be lower in the presence of an audit if $\Delta\phi_1\xi_1 \leq \Delta\phi_0\xi_0$. Other things equal, if the return on road quality is perceived to be smaller in the presence of an audit ($\xi_1 < \xi_0$) then participation will be lower.

[21]Note that if $\Delta\phi_1 = 0$ then $\Delta\phi_1\xi_1 \leq \Delta\phi_0\xi_0$ holds trivially regardless of what ξ_0 and ξ_1 are. Intuitively, in making your participation decision, it is trivial whether you can get a different quality road (differing ξ's mechanism) if your marginal participation will not change the likelihood that you get that road (non-pivotal participation mechanism)

labor differs across audit and non-audit villages with the same level of corruption. There is evidence of a higher share of gravel used in the presence of audit; however, when looking at all materials combined (column (5)), the overall difference is close to zero.

We also look at the infrastructure quality score assigned by auditors and engineers within audited villages. These quality scores are an index from a survey where auditors and engineers filled out a long checklist rating the infrastructure quality of the project as "satisfactory", "deficient", and "very deficient". Higher scores denote higher construction quality (potentially higher ξ in terms of our model).[22] Columns (7) and (8) compare these scores across monitoring intervention villages conditional on the same level of corruption.[23] We find no evidence that, conditional on the same level of malfeasance, the quality score assigned to these roads differs in the presence of a monitoring technology. Overall, the suggestive evidence presented here along with the likely presence of the non-pivotal participation mechanism suggest that the differing public benefits (differing ξ's) mechanism is unlikely to be a key channel explaining the drop in participation in the presence of audits.

In the case of the distribution of pro-social norms $F(.)$, participation could be lower in audit villages if norms, i.e., λ's, tend to be lower in audit villages. This could happen if the central government selects villages with low pro-social norms to be audited more often. In such cases, participation on average will be lower in audited villages due to endogenous selection by the auditing agency. Within the context of the paper, however, this is trivial as the auditing intervention was randomized.

5 Discussion and Additional Results

This section discusses and presents evidence on the relative effectiveness of a purely top-down *versus* a purely bottom-up monitoring strategy. It then compares the distortionary effects of grassroots interventions relative to centralized audits in terms of the substitution across

[22]Refer to Olken (2007) for more details on the quality scores.

[23]Since these scores are only reported for audited villages, we cannot estimate an "audit" effect.

different types of corrupt behavior that each may cause.

5.1 Purely Top-down *versus* Purely Bottom-up Strategy

An important policy question is whether a corruption deterring strategy should follow a purely top-down, or a purely bottom-up strategy, or a combination of both. Overall, this paper shows that despite policies promoting joint implementation of both interventions, top-down monitoring seems to depress rather than complement grassroots efforts. We proceed by presenting evidence on the relative effectiveness of a purely top-down (audit) strategy versus a purely bottom-up (invitations and comment forms) strategy.

We proceed by estimating the completely interacted version of equation (1) using alternative clustering and fixed effects that allow estimating an audit effect.[24] Specifically, we estimate:

$$\begin{aligned} Missing\ Expenditures_{ijk} &= \alpha_1 + \alpha_2 Audit_{jk} + \alpha_3 IC_{ijk} \times School_{ijk} + \alpha_4 IC_{ijk} \times (1 - School_{ijk}) \\ &+ \alpha_5 Audit_{jk} \times IC_{ijk} \times School_{ijk} + \alpha_6 Audit_{jk} \times IC_{ijk} \times (1 - School_{ijk}) + \delta_k + \epsilon_{ijk} \quad (6) \end{aligned}$$

where $Audit_{jk}$ equals one if a village in subdistrict j of district k had a government audit in place. δ_k denotes the audit stratum fixed effect (i.e., a district fixed effect). In another specification, we use an engineering team fixed effect. This enables identification of the coefficient on "Audit", α_2. The other terms are defined as in equation (1). Clustering is done at the subdistrict level given that audit randomization was done at this level. The purely audit effect is given by α_2, while the purely bottom-up (invitations and comment forms) effect is given by α_3.

Columns (5) and (6) of Table 3 present the estimates of α_2 and α_3 from Equation (6). We highlight an interesting result. A purely bottom-up strategy that is effectively implemented

[24]Recall that equation (1) used subdistrict fixed effects and clustering at the village level which were appropriate to estimate the effect of the participation interventions but not to estimate the effect of the audit (audit assignment was constant within subdistricts.

(e.g., elite capture is resolved) is as effective as a government audit in decreasing the percent of missing expenditures. Specifically, we find that in audit-only villages, missing expenses decreased by 7.8 percentage points while in invitations-only villages missing expenses decreased by 7.1 percentage points (column (5)). As shown in the row labeled "P-value ($\alpha_2 = \alpha_3$)" which provides the p-value for the Wald test comparing coefficients α_2 and α_3; the difference in the effects is not statistically significant suggesting that the two effects are statistically identical. Using engineering team fixed effects (column (6)) yields similar results.[25]

Given this finding, it is difficult to assess whether a policy of purely top-down or purely bottom-up monitoring should be followed as the bottom-up approach seems to yield similar results when properly implemented. From a policy perspective, however, grassroots interventions can be more cost-effective to implement as top-down alternatives require the hiring, training, and employment of professional auditors. A carefully implemented grassroots initiative avoids these costs and potentially enables a more efficient "unbundling" of monitoring tasks. For instance, trained auditors can be employed in more specialized auditing tasks since ordinary citizens can monitor the day-to-day actions of local officials.

Combining the results in this section with the cost-benefit figures in Olken (2007), we perform a simple back-of-the-envelope calculation on the net benefits of the grassroots intervention versus the audit. Since, according to our results in this section, the school-distributed intervention performs as well as the audit, then the calculation simply deals with the costs of implementation while leaving all the changes in benefits from the project and corruption rents unchanged from Olken (2007). Therefore, all of the accrued benefits of the grassroots intervention come from cost saving. With this in mind, the grassroots intervention produces a net benefit of about $714 (USD) per village.[26] This is almost three times the net benefit

[25] Results are similar using missing expenses in roads and ancillary projects as the outcome variable. See Appendix Table A6 for these results.

[26] We focus our comparison using the numbers for the equal-weighted net benefits in Table 13 in Olken (2007). We use the time cost of $31 for attending the meetings while monetary costs and the associated dead-weight loss is assumed to be zero in the case of the grassroots intervention. This yields a total cost of treatment of $31 which compared to the $468 in corruption rents and $1,213 in project benefits yields the $714 net benefit.

of the audit. We also note, that as discussed in Olken (2007), the monetary benefits of the audit can potentially be achieved with a lower intensity audit if for instance, a 50% audit intensity (instead of the experimental 100%) achieves the same level of corruption deterrence. In such case, the cost of the audit is much lower than the one used here for the comparison. The differences in net benefits between grassroots and audit will be smaller but unlikely to change the main takeaway, i.e., that at similar levels of corruption-deterring performance, a grassroots intervention is much more cost effective than an audit.[27]

Additionally, exposing communities to grassroots interventions can have positive externalities. Unlike audits which are context-specific and mostly external to the communities, exposure to grassroots interventions can help instill pro-social norms among individuals. This, in turn, can help establish a systematic monitoring presence in communities as potentially corrupt officials have a constant "threat" of monitoring by a more engaged community. Less formally, these interventions can help communities get used to participating in the accountability process. This can potentially have long-term effects although this is speculative and an interesting area to be further explored.

5.2 Substitution Across Corrupt Offenses

Olken (2007) finds evidence that in audit villages individuals related to village government officials or project heads are more likely to be employed in the project. This provides suggestive evidence that audits may lead officials to substitute from deterred theft to nepotism. We explore whether the presence of community monitoring also leads to similar substitution patterns towards nepotism. To do this, we replicate the results in Olken (2007) for non-audit villages and using the household survey collected as part of the project.[28] We estimate the

[27]For instance, again using Olken (2007) figures and assuming a 50% intensity audit carries half the costs of the 100% audit would yield an estimated net benefit of $480. The net benefits from the grassroots intervention are still 49 percent higher.

[28]The household survey was conducted towards the finish of the construction projects. The household survey was conducted using a stratified random sampling strategy. On average, each village surveyed contained between 6 and 13 respondents per village. Refer to Olken (2007) for more information on the design of the survey.

following equation for non-audit villages:

$$Worked_{ij} = \alpha_1 + \alpha_2 IC_{ij} + \alpha_3 Related_{hij} + \alpha_4 IC_{hij} \times Related_{hij} + \alpha_5 X_{hij} + \delta_j + \epsilon_{hij} \quad (7)$$

where $Worked_{hij}$ is whether respondent in household h in village i reports working in the project for pay. $Related_{hij}$ is whether the respondent is related to either a village government official or a project head. X_{hij} is a vector of household and individual-level controls.[29] j denotes subdistrict and δ_j indicates subdistrict fixed effects. We are interested in coefficient α_4 which indicates whether community meetings in non-audit villages also led to increased nepotism.

Table 7 presents the estimated coefficients from Equation (7). Column (1) looks at individuals related to village government officials, column (2) looks at individuals related to project heads, column (3) combines both. In all cases, the estimates of α_4 are negative and statistically insignificant. This suggests that the presence of grassroots monitoring in non-audit villages did not lead to higher employment of connected individuals in the projects. In contrast to the results from audits leading to nepotism, we find no convincing evidence that community monitoring had the same disruptive effect. We offer two possible explanations: First, the monitoring technologies are intrinsically different. Audits (top-down monitoring) are centralized and external to the community. Community monitoring, on the other hand, is "internal" and participatory in nature. Therefore, forms of malfeasance such as nepotism are more easily detected using community monitoring. Simply put, community monitors/citizens "know the ground" better than external auditors. Second, incentives to deter nepotism are higher for community monitors/citizens than for external auditors since ordinary citizens can be personally affected by the incidence of nepotism (i.e., forgone job opportunity).

[29]The controls included are age, sex, and years of education of respondent and number of social activities and average expenses in the household.

6 Conclusion

This paper proposes an alternative mechanism to explain why bottom-up monitoring strategies seem to under-perform relative to top-down strategies. We argue that in a setting where both bottom-up and top-down strategies are implemented, effective top-down alternatives can actually lead to a crowding-out of individuals' participation in the monitoring process. In other words, the presence of effective and credible top-down monitoring can actually undermine the participation goals of a competing grassroots intervention.

Building on Olken (2007) Indonesia corruption experiment, we find considerable evidence that the participation intervention was successful in villages where this was the only monitoring intervention. However, in villages where a credible government audit was being simultaneously implemented, the effect of the grassroots intervention was close to zero. We provide further evidence that the contrasting effect is the result of the government audit hindering participation and engagement in the accountability meetings. After carefully analyzing data on project-related accountability meetings, we find that in villages where an audit is in place, individuals are less likely to attend and participate during the accountability meetings. They are also significantly less motivated to voice general problems, corruption-related problems, and to take serious actions to address these problems. This offers a stark contrast to non-audited villages where the voicing of problems, and proactive participation was more significant.

Encouragingly, we provide evidence that, when properly implemented, a grassroots intervention can be: (i) as effective as a top-down monitoring strategy in reducing corruption, (ii) almost three times as cost-effective as a comparable audit, and (iii) lead to less distortions in the form of dishonest officials substituting across different types of corrupt behavior.

References

Banerjee, A., R. Hanna, J. Kyle, B. A. Olken, and S. Sumarto (2018). Tangible information and citizen empowerment: Identification cards and food subsidy programs in indonesia. *Journal of Political Economy 126*(2), 451–491.

Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in india. *American Economic Journal: Economic Policy 2*(1), 1–30.

Barr, A., F. Mugisha, P. Serneels, and A. Zeitlin (2012). Information and collective action in community-based monitoring of schools: Field and lab experimental evidence from uganda. *Unpubl. Pap.*

Bjorkman, M. and J. Svensson (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in uganda. *Quarterly Journal of Economics*.

Bowles, S. and S. Polania-Reyes (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature 50*(2), 368–425.

Callen, M. and J. D. Long (2015). Institutional corruption and election fraud: Evidence from a field experiment in afghanistan. *American Economic Review 105*(1), 354–381.

Chassang, S. and G. P. i. Miquel (2018). Crime, intimidation, and whistleblowing: A theory of inference from unverifiable reports. *Review of Economic Studies*.

Ferraz, C. and F. Finan (2008). Exposing corrupt politicians: The effects of brazil's publically released audits on electoral outcomes. *Quarterly Journal of Economics*.

Gneezy, U., S. Meier, and P. Rey-Biel (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives 25*(4), 191–210.

Gonzalez, R. (2020). Cellphone access and election fraud: Evidence from a spatial regression discontinuity design in afghanistan. *Forthcoming at American Economic Journal: Applied Economics*.

Lowes, S., N. Nunn, J. A. Robinson, and J. L. Weigel (2017). The evolution of culture and institutions: Evidence from the kuba kingdom. *Econometrica 85*(4), 1065–1091.

Olken, B. A. (2007). Monitoring corruption: evidence from a field experiment in indonesia. *Journal of political Economy 115*(2), 200–249.

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics 93*(7-8), 950–964.

Pandey, P., S. Goyal, and V. Sundararaman (2009). Community participation in public schools: impact of information campaigns in three indian states. *Education economics 17*(3), 355–375.

Pradhan, M., D. Suryadarma, A. Beatty, M. Wong, and A. Gaduh (2014). Artha, 2014. improving educational quality through enhanced community participation: Results from a randomised field experiment in indonesia. *American Economic Journal: Applied Economics 6*(2), 105–126.

Raffler, P., D. N. Posner, and D. Parkerson (2019). The weakness of bottom-up accountability: Experimental evidence from the ugandan health sector. *Working paper*.

Serra, D. (2011). Combining top-down and bottom-up accountability: evidence from a bribery experiment. *The journal of Law, Economics, & Organization 28*(3), 569–587.

Tabellini, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics 123*(3), 905–950.

Tables

Table 1: Summary Statistics by Bottom-Up Monitoring Status

		Invitations + Comments		Differences			
	Control	Schools	Heads	(2)-(1)	(3)-(1)	(2)-(3)	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Village population (000s)	4.225	4.810	4.118	0.585	-0.107	0.692	565
	(0.206)	(0.253)	(0.196)	(0.221)***	(0.211)	(0.257)***	
Mosques per 1,000	1.474	1.374	1.449	-0.100	-0.025	-0.074	565
	(0.066)	(0.077)	(0.071)	(0.076)	(0.081)	(0.086)	
Total budget (Rp. millions)	81.983	81.285	79.194	-0.698	-2.789	2.091	565
	(3.423)	(5.577)	(3.142)	(6.014)	(3.472)	(5.905)	
Number subprojects	2.757	2.702	2.798	-0.055	0.041	-0.096	554
	(0.103)	(0.108)	(0.115)	(0.118)	(0.111)	(0.137)	
Percent households poor	0.405	0.408	0.419	0.003	0.014	-0.011	560
	(0.016)	(0.016)	(0.018)	(0.021)	(0.018)	(0.023)	
Distance to subdistrict	5.464	5.321	5.135	-0.143	-0.329	0.186	565
	(0.287)	(0.381)	(0.315)	(0.449)	(0.391)	(0.472)	
Village head education	11.503	11.648	11.228	0.146	-0.275	0.421	562
	(0.211)	(0.201)	(0.220)	(0.288)	(0.268)	(0.310)	
Village head age	42.848	43.489	44.312	0.641	1.464	-0.823	562
	(0.578)	(0.621)	(0.572)	(0.843)	(0.828)*	(0.854)	
Village head salary	2.652	3.580	3.055	0.927	0.403	0.524	559
	(0.333)	(0.495)	(0.394)	(0.509)*	(0.364)	(0.502)	
Mountainous dummy	0.361	0.375	0.367	0.014	0.006	0.008	563
P-value (joint significance)				0.468	0.817	0.649	

Note—“Invitations + Comments” refers to both “Invitations” and “Invitations and Comment forms” treatments combined. “Schools” and “Heads” refers to whether the invitations and comment forms treatment was distributed via schools or neighborhood heads, respectively. Columns (4)-(6) provide the mean comparison tests between the indicated columns. Standard errors clustered at the subdistrict level. *, **, *** denotes significant at 10, 5, and 1 percent, respectively. “P-value (joint significance)” refers to the p-value from the joint significance test of all coefficients obtained from a Probit model using all controls to predict school distribution restricting sample to school distribution and control villages (column (4)), neighborhood head distribution restricting sample to neighborhood head distribution and control villages (column (5)), and school distribution restricting sample to school or head distribution (column (6)).

Table 2: Summary Statistics of Missing Expenses by All Treatment Categories

		Invitations + Comments		Differences	
	Control	Schools	Heads	(2)-(1)	(3)-(1)
	(1)	(2)	(3)	(4)	(5)
Control/No audit ($N = 253$)	0.303	0.212	0.305	-0.090*	0.002
	(0.039)	(0.051)	(0.044)	(0.052)	(0.042)
Audit ($N = 224$)	0.192	0.201	0.182	0.009	-0.010
	(0.041)	(0.041)	(0.043)	(0.056)	(0.044)
Observations ($N = 477$)	162	151	164		

Note— "Invitations + Comments" refers to both "Invitations" and "Invitations and Comment forms" treatments combined. Randomization of audits was independent of the randomization into invitations or invitations plus comment forms. Refer to section 2.1 in the text or Olken (2007) for more details on the experimental design. "Schools" and "Heads" refers to whether the invitations and comment forms treatment was distributed via schools or neighborhood heads, respectively. Standard errors presented in parenthesis and clustered at the subdistrict level. Columns (4) and (5) provide the mean comparison tests between the indicated columns.

Table 3: Effect of Grassroots Monitoring on Missing Expenses

	Dependent variable: Percent Missing Expenses					
	All	No Audit	Audit	Full	Full	Full
	(1)	(2)	(3)	(4)	(5)	(6)
Invitations and Comments	-0.029	-0.083**	0.024	-0.083**	-0.071*	-0.093**
	(0.032)	(0.039)	(0.050)	(0.039)	(0.039)	(0.046)
Audit					-0.078	-0.100**
					(0.049)	(0.050)
Invitations and Comments × Audit				0.107*	0.070	0.084
				(0.063)	(0.066)	(0.069)
Mean	0.252	0.303	0.192	0.252	0.252	0.252
Observations	477	253	224	477	477	477
P-value ($\alpha_2 = \alpha_3$)	-	-	-	-	0.892	0.911
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Engineer FE	No	No	No	No	No	Yes

Note— Columns (1)-(4) present results from estimating Equation (1). Invitations and Comments distribution via schools. Columns (5) and (6) present results from estimating Equation (6). "P-value ($\alpha_2 = \alpha_3$)" refers to the p-value from test of equality between coefficients on "Invitations and Comments" and "Audit" in Equation (6) and presented in Columns (5) and (6). Stratum (subdistrict) fixed effects are included in columns (1)-(4). Stratum (district) fixed effects in column (5). Engineer fixed effects are included in column (6). Standard errors clustered at the village level in columns (1)-(4) and at the subdistrict level in columns (5) and (6). *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Table 4: Effect on Attendance and Active Participation by Audit Status

	Attendance of Nonelite			Number of Nonelite who talk		
	All (1)	No Audit (2)	Audit (3)	All (4)	No Audit (5)	Audit (6)
Invitations and Comments	11.861***	12.967***	10.524***	0.253***	0.344***	0.143
	(1.073)	(1.591)	(1.380)	(0.064)	(0.091)	(0.090)
Mean	24.153	23.860	24.496	0.944	0.881	1.018
Observations	1,775	956	819	1,775	956	819
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Meeting FE	Yes	Yes	Yes	Yes	Yes	Yes
P-value (No audit=Audit)		0.244			0.117	

Note— Results come from estimating eq. (2) with the participation variables shown in the first row. "P-value (No audit=Audit)" refers to the p-value on the interaction between audit and the treatment of comments and invitations. Each observation is a single village meeting. "Nonelite" refers to individuals that have no official position in the village or the project (Olken, 2007). Stratum (subdistrict) fixed effects are included; since audit is constant within a subdistrict, the audit variable is automatically captured by the stratum fixed effect. Robust standard errors are in parentheses, adjusted for clustering at the village level. *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Table 5: Effect on Meetings by Audit Status

	Number of Problems			Any Corruption Related Problem			Serious Response Taken		
	All (1)	No Audit (2)	Audit (3)	All (4)	No Audit (5)	Audit (6)	All (7)	No Audit (8)	Audit (9)
Invitations and Comments	0.088	0.188**	-0.033	0.026**	0.027*	0.026	0.006	0.021**	-0.013
	(0.058)	(0.078)	(0.087)	(0.011)	(0.015)	(0.016)	(0.007)	(0.010)	(0.009)
Mean	1.177	1.153	1.205	0.06	0.075	0.054	0.026	0.037	0.012
Observations	1,783	963	820	1,783	963	820	1,783	963	820
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Meeting FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P-value (No audit=Audit)		0.057			0.945			0.013	

Note— Results come from estimating eq. (2) with the participation variables shown in the first row. "P-value (No audit=Audit)" refers to the p-value on the interaction between audit and the treatment of comments and invitations. Each observation represents one meeting. "Serious response" is defined as "agreeing to replace a supplier or village office, agreeing that money should be returned, agreeing to an internal village investigation, asking for help from district project officials, or requesting an external audit" (Olken, 2007). Robust standard errors are in parentheses, adjusted for clustering by village. *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Table 6: Differences in Project Characteristics between Audit and non-Audit villages

		Share of road expenses in:					Infrastructure Quality score	
	Project size (USD)	Sand	Rock	Gravel	All Materials	Unskilled labor	Auditors	Engineers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Invitations and Comments	-0.043	-0.009	0.018	-0.015	0.002	-0.002	-0.087	-0.096
	(0.061)	(0.008)	(0.013)	(0.012)	(0.012)	(0.012)	(0.091)	(0.091)
Audit	-0.101	-0.011	0.022	-0.029*	0.005	-0.005		
	(0.078)	(0.010)	(0.019)	(0.017)	(0.017)	(0.017)		
Invitations and Comments × Audit	0.097	-0.000	-0.009	0.062**	0.001	-0.001		
	(0.086)	(0.012)	(0.021)	(0.031)	(0.019)	(0.019)		
Missing expenditures	0.077	0.007	0.012	0.033	-0.036	0.036	0.048	0.018
	(0.078)	(0.010)	(0.026)	(0.060)	(0.023)	(0.023)	(0.151)	(0.135)
Mean	8.99	0.10	0.48	0.12	0.80	0.20	-0.02	0.04
Observations	476	477	477	477	477	477	219	212

Note— Columns (1)-(6) include district fixed effects and standard errors clustered at the subdistrict level. Columns (7), (8) include subdistrict fixed effects and standard errors clustered at the village level. Project size is in logs. Quality score refers to the score given by auditors and engineers, respectively, on the project infrastructure quality. Auditors and engineers filled out the same checklist rating the infrastructure quality of the project as satisfactory, deficient, very deficient. Scores are normalized to have mean zero and standard deviation one. See Olken (2007) for more detail. . *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

TABLE 7: Effect of Grassroots Monitoring on Nepotism

	Dependent variable: Individual Worked in Project		
	(1)	(2)	(3)
Invitations and Comments	-0.034	-0.044	-0.049
	(0.032)	(0.029)	(0.046)
Village Gov't Family Member	0.015		0.014
	(0.041)		(0.041)
Invitations and Comments × Village Gov't Family Member	-0.035		-0.035
	(0.050)		(0.050)
Project Head Family Member		-0.015	-0.027
		(0.082)	(0.083)
Invitations and Comments × Project Head Family Member		-0.031	-0.017
		(0.090)	(0.091)
Mean	0.300	0.300	0.300
Observations	1789	1789	1789

NOTE— Results come from estimating Equation (7). Results include subdistrict fixed effects. Standard errors clustered at the village level. All specifications include controls for age, sex, and years of education of respondent and number of social activities and average expenses in the household. *, **, *** denotes significant at 10, 5, and 1 percent, respectively

Appendix A Additional Tables

Table A1: Effect of Grassroots Monitoring on Missing Expenses, Engineering Team FE

	Dependent variable: Percent Missing Expenses					
	All (1)	No Audit (2)	Audit (3)	All (4)	No Audit (5)	Audit (6)
Invitations and Comments	-0.029	-0.083**	0.024	-0.052	-0.095**	0.001
	(0.032)	(0.039)	(0.050)	(0.036)	(0.042)	(0.053)
Mean	0.252	0.303	0.192	0.252	0.303	0.192
Observations	477	253	224	477	253	224
Stratum FE	Yes	Yes	Yes	No	No	No
Engineer FE	No	No	No	Yes	Yes	Yes
P-value (No audit=Audit)		0.09			0.24	

Note— Results come from estimating equation (1). "P-value (No audit=Audit)" refers to the p-value on the interaction term of audit treatment and the treatment of comments and invitations. Stratum (subdistrict) fixed effects are included in the first three estimations and engineer fixed effects are included in the last three estimations. Standard errors clustered at the village level in columns (1)-(3) and at the subdistrict level in columns (4)-(6). *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Table A2: Effect of Grassroots Monitoring on Missing Expenses, All Distribution Methods

	Dependent variable: Percent Missing Expenses					
	All	No Audit	Audit	All	No Audit	Audit
	(1)	(2)	(3)	(4)	(5)	(6)
Invitations and Comments	-0.021	-0.041	0.003	-0.028	-0.047	0.001
	(0.027)	(0.033)	(0.045)	(0.027)	(0.032)	(0.043)
Mean	0.252	0.303	0.192	0.252	0.303	0.192
Observations	477	253	224	477	253	224
Stratum FE	Yes	Yes	Yes	No	No	No
Engineer FE	No	No	No	Yes	Yes	Yes
P-value (No audit=Audit)		0.44			0.50	

Note— Results come from estimating version of equation (1) that does not separate by school distribution (i.e., $Y_{ij} = \alpha_1 + \alpha_2 IC_{ij} + \delta_j + \epsilon_{ij}$, where all terms are defined as in equation (1) in the text). "P-value (No audit=Audit)" refers to the p-value on the interaction term of audit treatment and the treatment of comments and invitations. Stratum (subdistrict) fixed effects are included in the first three estimations and engineer fixed effects are included in the last three estimations. Standard errors clustered at the village level in columns (1)-(3) and at the subdistrict level in columns (4)-(6). *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Table A3: Effect of Grassroots Monitoring on Missing Expenses, Alternative Outcome

	Dependent variable: Percent Missing Expenses in Roads and Ancillary Projects					
	All	No Audit	Audit	All	No Audit	Audit
	(1)	(2)	(3)	(4)	(5)	(6)
Invitations and Comments	-0.067**	-0.090*	-0.046	-0.069*	-0.106**	-0.036
	(0.032)	(0.046)	(0.044)	(0.035)	(0.044)	(0.049)
Mean	0.247	0.291	0.199	0.247	0.291	0.199
Observations	538	281	257	538	281	257
Stratum FE	Yes	Yes	Yes	No	No	No
Engineer FE	No	No	No	Yes	Yes	Yes
P-value (No audit=Audit)		0.49			0.50	

Note— Results come from estimating equation (1). "P-value (No audit=Audit)" refers to the p-value on the interaction term of audit treatment and the treatment of comments and invitations. Stratum (subdistrict) fixed effects are included in the first three estimations and engineer fixed effects are included in the last three estimations. Standard errors clustered at the village level in columns (1)-(3) and at the subdistrict level in columns (4)-(6). *, **, *** denotes significant at 10, 5, and 1 percent, respectively

Table A4: Effect on Attendance and Active Participation by Audit Status, Distribution via Schools

	Attendance of Nonelite			Number of Nonelite who talk		
	All	No Audit	Audit	All	No Audit	Audit
	(1)	(2)	(3)	(4)	(5)	(6)
Invitations and Comments	11.972***	12.358***	11.478***	0.231**	0.361**	0.086
	(1.476)	(2.241)	(1.881)	(0.097)	(0.160)	(0.106)
Mean	24.331	27.358	20.353	0.966	1.045	0.863
Observations	1,657	889	768	1,657	889	768
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Meeting FE	Yes	Yes	Yes	Yes	Yes	Yes
P-value (No audit=Audit)		0.76			0.15	

Note— Results come from estimating equation (1) with the dependent variables the participation variables shown in the first row. The p-values are calculated from regressions of the dependent variables on the interaction of not being audited and the treatment of comments and invitations. Data are taken from the meeting survey. The results are estimated for all villages, villages that were audited and villages that were not audited. Each observation is a single village meeting. Stratum (subdistrict) fixed effects are included; since audit is constant within a subdistrict, the audit variable is automatically captured by the stratum fixed effect. Robust standard errors are in parentheses, adjusted for clustering at the village level. *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Table A5: Effect on Meetings by Audit Status, Distribution via Schools

	Number of Problems			Any Corruption Related Problem			Serious Response Taken		
	All	No Audit	Audit	All	No Audit	Audit	All	No Audit	Audit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Invitations and Comments	-0.045	0.028	-0.135	0.035**	0.056**	0.012	0.000	0.017	-0.020*
	(0.069)	(0.094)	(0.102)	(0.015)	(0.022)	(0.018)	(0.009)	(0.014)	(0.011)
Mean	1.172	1.160	1.186	0.062	0.075	0.047	0.026	0.038	0.013
Observations	1,665	896	769	1,665	896	769	1,665	896	769
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Meeting FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P-value (No audit=Audit)		0.24			0.13			0.03	

Note— Results come from estimating equation (1), with the dependent variables the outcome of meetings shown in the first row. The p-values are calculated from regressions of the dependent variables on the interaction of not being audited and the treatment of comments and invitations. Data are taken from the meeting survey. The results are estimated for all villages, villages that were audited and villages that were not audited. Each observation represents one village. “Serious response” is defined as agreeing to replace a supplier or village office, agreeing that money should be returned, agreeing to an internal village investigation, asking for help from district project officials, or requesting an external audit. Estimation is by OLS. Robust standard errors are in parentheses, adjusted for clustering by village. *, **, *** denotes significant at 10, 5, and 1 percent, respectively.“RI P-value” refers to the p-value calculated using randomization inference.

Table A6: Audit *versus* Grassroots, Alternative Outcome

	Dependent variable: Percent Missing Expenses in Roads and Ancillary Projects		
	(1)	(2)	(3)
Audit	-0.094*	-0.093*	-0.097*
	(0.050)	(0.052)	(0.055)
Invitations and Comments	-0.075*	-0.088*	-0.087*
	(0.044)	(0.048)	(0.051)
P-value ($\alpha_2 = \alpha_3$)	0.711	0.931	0.876
Observations	538	538	538
Mean dependent variable	0.25	0.25	0.25
Stratum FE	Yes	No	No
Engineer FE	No	Yes	No

Note— "P-value ($\alpha_2 = \alpha_3$)" refers to the p-value from test of equality between coefficients on "Invitations and Comments" and "Audit". Column (1) uses audit stratum (district) fixed effects. Column (2) uses engineering team fixed effects. Column (3) uses no fixed effects. All specification use standard errors clustered at the subdistrict level (level of randomization for audit treatment). *, **, *** denotes significant at 10, 5, and 1 percent, respectively.

Appendix B Model Extensions

Assume that the individual is an expected utility maximizer. When corruption is deterred the individual receives public gain $\xi > 0$, this can be thought of as the utility gain from a higher quality road or public good. Participation carries a cost $c > 0$ but can bring private utility gain $\lambda \geq 0$ which can be interpreted as "warm glow" received from engaging in pro-social behavior. Assume that the individual can choose his degree of participation $p \in [0, 1]$ with p closer to 1 meaning full participation and engagement. In deciding how much to participate, the individual assesses that the probability that corruption will be deterred conditional on his degree of participation p and audit intensity $a \in [0, 1]$ is given by $\phi_{a,p}$. The optimization problem of the individual is thus given by:

$$\max_{p} \quad \phi_{a,p}[\xi + \lambda(p) - c(p)] + (1 - \phi_{a,p})[\lambda(p) - c(p)] \tag{B1}$$

where it is assumed that the participation costs and private gains from participation depend on the degree of participation p. Furthermore, assume that benefits are concave and increasing in participation ($\lambda'(p) > 0$, $\lambda''(p) < 0$) and costs are convex and increasing in participation ($c'(p) > 0$, and $c''(p) > 0$). Differentiating (B1), one obtains first order condition:

$$\lambda'(p) - c'(p) + \xi \frac{\partial \phi_{a,p}}{\partial p} = 0 \tag{B2}$$

Applying the implicit function theorem on (B2), one obtains that:

$$\frac{dp*}{da} = -\frac{\xi \frac{\partial \phi_{a,p}}{\partial p \partial a}}{\lambda''(p) - c''(p) + \xi \frac{\partial \phi^2_{a,p}}{\partial p^2}} \tag{B3}$$

Therefore, the effect of audit intensity on the equilibrium level of participation will be negative as long as $\frac{\partial \phi_{a,p}}{\partial p \partial a} < 0$.[30] In essence, this conclusion is actually the continuous version of

[30] Plausible assuming that participation has diminishing returns on the likelihood of deterrence (e.g., the return to an additional corruption complaint becomes smaller after a number of similar complaints have

the condition described in the main text, i.e., that the individual lowers his participation if: $\Delta\phi_1 - \Delta\phi_0 = (\phi_{1,1} - \phi_{1,0}) - (\phi_{0,1} - \phi_{0,0}) < 0$. This implies that participation will decrease if the individual believes that higher audit intensity lowers the marginal contribution of his participation.

already been recorded), then the denominator in (B3) is negative.