MONITORING CORRUPTION:
EVIDENCE FROM A FIELD EXPERIMENT IN INDONESIA

Benjamin A. Olken

**ABSTRACT**

This paper uses a randomized field experiment to examine several approaches to reducing corruption. I measure missing expenditures in over 600 village road projects in Indonesia by having engineers independently estimate the prices and quantities of all inputs used in each road, and then comparing these estimates to villages' official expenditure reports. I find that announcing an increased probability of a government audit, from a baseline of 4 percent to 100 percent, reduced missing expenditures by about 8 percentage points, more than enough to make these audits cost-effective. By contrast, I find that increasing grass-roots participation in the monitoring process only reduced missing wages, with no effect on missing materials expenditures. Since materials account for three-quarters of total expenditures, increasing grass-roots participation had little impact overall. The findings suggest that grass-roots monitoring may be subject to free-rider problems. Overall, the results suggest that traditional top-down monitoring can play an important role in reducing corruption, even in a highly corrupt environment.

Benjamin A. Olken
Harvard Society of Fellows
78 Mount Auburn St.
Cambridge, Massachusetts 02138
and NBER
bolken@nber.org

## 1. Introduction

Corruption is a significant problem in much of the developing world. In many cases, corruption acts like a tax, adding to the cost of providing public services and conducting business. Often, though, the efficiency costs of corruption can be far worse.[1] Indeed, it has been suggested that corruption may be a major contributor to the low growth rates of many developing countries (Mauro 1995).

Despite the importance of the problem, the inherent difficulty of directly measuring corrupt activity has meant that there is relatively little evidence, and therefore relatively little consensus, on how to best reduce corruption. One approach to reducing corruption, dating back at least to Becker and Stigler (1974), suggests that the right combination of monitoring and punishments can control corruption. In practice, however, the very individuals tasked with monitoring and enforcing punishments may themselves be corruptible. In that case, increasing the probability that a low-level official is monitored by a higher-level official could result only in a transfer between the officials, not in a reduction of corruption.[2] Whether one can actually control corruption by increasing top-down monitoring in such an environment is an open, and important, empirical question.

An alternative approach to reducing corruption, which has gained prominence in recent years, is to increase grass-roots participation by community members in local-level monitoring. Community participation is now regarded in much of the development community as the key not only to reduced corruption, but to improved public service delivery more generally. For example, the entire 2004 World Development Report is devoted to the idea of "putting poor people at the

---

[1] See, for example, Krueger (1974) and Shleifer and Vishny (1993) for examples of how the efficiency costs of corruption can substantially exceed the amount stolen itself.

[2] Cadot (1987), for example, discusses this possibility, and shows that this type of multi-tiered corruption can lead to multiple equilibria in corruption.

center of service provision: enabling them to monitor and discipline service providers, amplifying their voice in policymaking, and strengthening the incentives for service providers to serve the poor." (World Bank 2004) The idea behind the grass-roots approach is that community members are the people who benefit from a successful program, and so may have better incentives to monitor than disinterested central government bureaucrats (Stiglitz 2002). Of course, this approach has potential drawbacks as well – for example, monitoring public projects is a public good, so there may be a serious free rider problem. Grass-roots monitoring may also be prone to capture by local elites (Bardhan 2002, Bardhan and Mookherjee forthcoming). Given these countervailing forces, whether grass-roots monitoring can actually succeed in reducing corruption is also an empirical question.[3]

To examine these alternative approaches to fighting corruption, I designed and conducted a randomized, controlled field experiment in 608 Indonesian villages. At the time the study started, each village in the study was about to start building a village road as part of a nationwide village-level infrastructure project. To examine the impact of external monitoring, I randomly selected some villages to be told, after funds had been awarded but before construction began, that their project would subsequently be audited by the central government audit agency. This amounted to increasing the probability of an external government audit in those villages from a baseline of about 4 percent to essentially 100 percent. Government audits carry with them the theoretical possibility of criminal action, though this is quite rare; more important, the results of the audits were read publicly to an open village meeting by the auditors, and so could result in substantial social sanctions. The audits were subsequently conducted as promised.

---

[3] Several authors have found suggestive evidence in both micro and macro cross-sectional data that higher levels of "voice" are associated with lower levels of corruption. Rose-Ackerman (2004) provides a summary of much of the work on this topic to date.

To investigate the impact of increasing community participation in the monitoring process, I designed two different experiments that sought to increase grass-roots monitoring in the project. Specifically, the experiments sought to enhance participation at "accountability meetings," the village-level meetings where project officials account for how they spent project funds. In one experiment, hundreds of invitations to these meetings were distributed throughout the village, to encourage direct participation in the monitoring process and to reduce elite dominance of the process. In the second experiment, an anonymous comment form was distributed along with the invitations, providing villagers an opportunity to relay information about the project without fear of retaliation. This comment form was then collected before the accountability meetings in sealed drop-boxes, and the results were summarized at the meetings. Both of these experimental interventions were successful in raising grass-roots participation levels—the invitations increased the number of people participating in the accountability meetings by about 40 percent, and the comment forms generated hundreds of comments about the project, both good and bad, in each village.

To evaluate the impact of these experiments on corruption, one needs a measure of corruption. Traditionally, much of the empirical work on corruption has been based on perceptions of corruption, rather than on direct measures of corruption.[4] This paper, however, builds on a small but growing literature that examines corruption by comparing two measures of the same quantity, one "before" and one "after" corruption has taken place.[5] To do this in the

---

[4] The use of perceptions-based measures of corruption in economics was pioneered by Mauro (1995) and forms the basis of the much-cited Transparency International Corruption Index (Lambsdorff 2003). More recent work using perceptions-based measures is summarized in Rose-Ackerman (2004).
[5] For example, Reinikka and Svensson (2004) examine corruption in educational expenditures, Fisman and Wei (2004) and Yang (2004) examine corruption in international trade, Di Tella and Schargrodsky (2003) examine corruption in hospital procurement, and Olken (forthcoming) examines theft from a government redistribution program. This paper differs slightly from much of this literature by comparing government reports to an independently constructed estimate, rather than comparing government reports to the reports of a different

context of the road projects, I assembled a team of engineers and surveyors who, after the projects were completed, dug core samples in each road to estimate the quantity of materials used, surveyed local suppliers to estimate prices, and interviewed villagers to determine the wages paid on the project. From these data, I construct an independent estimate of the amount each project actually cost to build, and then compare this estimate with what the village reported it spent on the project on a line-item by line-item basis. The difference between what the village claimed the road cost to build, and what the engineers estimated it actually cost to build, is the key measure of missing expenditures I examine in this paper. Since the village must account for every Rupiah it received from the central government, stolen funds must show up somewhere in the difference between reported expenditures and estimated actual expenditures.[6]

Using these data, I find that there were substantial reductions in missing expenditures associated with the audit experiment. In particular, I estimate that the audit treatment – i.e., increasing the probability of an audit from a baseline of 4 percent to 100 percent – was associated with reductions in missing expenditures of an average of 8 percentage points. These reductions came both from reductions in unaccounted-for materials procured for the project and in unaccounted-for labor expenditures. The audits were particularly effective in villages where the village head was up for reelection, and where the previous election results suggested that the upcoming election would be close, which suggests that external monitoring and democratic accountability may be complements, and more generally, that the interaction of the monitoring probability and the punishment level conditional on being caught is positive. Interestingly, I find

government agency or to a household survey. In that sense, it is related to Hsieh and Moretti (2005), who compare prices received by Iraq under the UN Oil-for-Food program to the world oil price.
[6] These 'missing expenditures', i.e., the difference between reported expenditures and my estimate of actual expenditures, may also include sources of losses other than pure theft. I discuss below how I constructed several test roads to estimate the typical amount of materials lost during construction, which I use to calibrate the missing expenditures measure to be 0 in a road where I know *a priori* that there was no corruption. I also discuss below how I use independent measures of the quality of road construction that are likely to be unrelated to corruption to differentiate between overall changes in the competence of road builders and corruption per se.

that the number of project jobs given to family members of project officials actually increased in response to the audits, which provides suggestive evidence that alternative forms of corruption may be substitutes. I compare the costs and benefits from the audits and find that on net the audits were highly cost effective.

By contrast, I find that the participation experiments – the invitations and the comment forms – were associated with much smaller, and statistically insignificant, reductions in overall missing expenditures. As discussed above, the interventions did raise community participation in the monitoring process. Moreover, villages in the invitations treatment were more likely to openly discuss corruption problems at the accountability meetings, and villages receiving both invitations and comment forms were more likely to take serious action at the meeting to resolve corruption-related problems. However, the magnitude of these changes in behavior at the meetings was small, and these treatments did not measurably reduce overall missing expenditures.

The small overall effects of the participation experiments on overall missing expenditures, however, mask substantial differences across types of expenditures. In particular, the invitations treatment substantially reduced missing labor expenditures, but had no effect whatsoever on missing materials expenditures. Since materials account for about three-quarters of total expenditures, the average impact was nonetheless small and statistically insignificant.

There are two potential explanations for why increasing grass-roots participation would have impacted labor expenditures but not materials – either community members had better information about labor expenditures than about materials, or community members had a greater incentive to monitor wage payments than materials payments (since they were the ones working on the project and would be the beneficiaries if wage payments had been stolen). In fact, the

results suggest that villagers had less information about missing labor expenditures than about missing materials expenditure. Furthermore, the invitations treatment was most effective when the workers came from inside the village and participated in monitoring. Combined, this suggests that grass-roots monitoring may be most effective when a subset of people personally stand to gain from reducing corruption, and therefore have strong incentives to monitor actively. Grass-roots monitoring may thus be subject to a free-rider problem.

The remainder of the paper is organized as follows. Section 2 discusses the setting in which the study takes place. Section 3 describes the experimental interventions. Section 4 describes the data used in the study. Section 5 presents the results of the experiments. Section 6 performs a cost-benefit analysis. Section 7 concludes.

## 2. Setting

The *Kecamatan* (Subdistrict) *Development Project*, or KDP, is a national government program, funded through a loan from the World Bank. KDP finances projects in approximately 15,000 villages throughout Indonesia each year. The data in this paper come from KDP projects in 608 villages in two of Indonesia's most populous provinces, East Java and Central Java, and were collected between September 2003 and August 2004.

In KDP, participating subdistricts, which typically contain between 10 and 20 villages, receive an annual block grant for three consecutive years. Every year, each village in the subdistrict makes a proposal for any combination of small-scale infrastructure and seed capital for microcredit cooperatives. The majority of villages (72%) propose an infrastructure project, plus a small amount for savings-and-loans for women. An inter-village forum ranks all of proposals according to a number of criteria, such as number of beneficiaries and project cost, and projects are funded according to the rank list until all funds have been exhausted.

If the project is funded, a meeting is held to plan construction, after which an elected implementation team procures materials, hires labor, and builds the project. The members of the implementation team receive an honorarium, limited in total to a maximum of 3 percent of the total cost of the project.[7] No contractors are used in construction.

A typical funded village receives funds on the order of Rp. 80 million (US$8,800) for infrastructure; these funds are often supplemented by voluntary contributions from village residents, primarily in the form of unpaid labor. These projects are large relative to ordinary local government activities; in 2001, the average village budget was only Rp. 71 million (US$7,800), so receiving a KDP project more than doubles average local government expenditures. The allocation to the village is lump-sum, so that the village is the residual claimant. In particular, surplus funds can be used, with the approval of a village meeting, for additional development projects, rather than having to be returned to the KDP program.

By far the most common type of infrastructure project proposed by villages is the surfacing of an existing dirt road with a surface made of sand, rocks, and gravel. These roads range in length from 0.5 – 3 km, and may either run within the village or run from the village to the fields. Dirt roads in Java are typically impassible during the rainy season; surfacing these roads allows them to be used year-round. To facilitate comparisons, the sample of villages considered in this paper is limited to villages with such non-asphalt road projects.

The project includes several mechanisms to ensure proper use of project funds. The primary mechanism is a series of village-level accountability meetings. Funds are released to the implementation team in three tranches, of 40%, 40%, and 20% of the funds, respectively. In order to obtain the second and third tranches of funds, the implementation team is required to

---

[7] In the data, the total honorarium payments to each member of the implementation team averaged Rp. 460,000, or approximately Rp. 150,000 ($17) per month of the construction. By comparison, median per-capita monthly expenditure in comparable areas in East and Central Java in 2003 was about Rp 140,000 ($15).

present an accountability report to an open village meeting, explaining how all funds were used. Only after that meeting has approved the accountability report is the next tranche of funds released. Similarly, in order to participate in the subsequent year of KDP, villages are required to present a final, cumulative accountability report at the end of the project, which similarly must be approved by a village meeting. Though open to the entire village, these meetings are typically attended by only 30-50 people, most of whom are members of the village elite, out of a total adult population of about 2,500.

In addition, subdistrict, district and provincial level project managers, engineers, and facilitators conduct overall supervision of all projects, and there is a provincial complaints-handling unit which investigates allegations of improprieties. Furthermore, each year, the project is audited by the independent government development audit agency, *Badan Pengawasan Keuangan dan Pembangunan* (BPKP). Each village-level project in the study area has about a 4% baseline chance of being audited by BPKP. Auditors come to the village, cross-check all of the financial records looking for irregularities, and inspect the physical infrastructure. Findings from the audits are sent to project officials for follow-up, and can potentially lead to criminal action, though prosecutions for village-level officials are rare in practice (Woodhouse 2004). More often, officials found to have stolen money are forced to publicly return the money, which can result in substantial social sanctions.

Corruption at the village level can occur in several ways.[8] First, implementation teams, potentially working with the village head, may collude with suppliers. Suppliers can inflate either the prices or the quantities listed on the official receipts to generate money for a kick-back to village and project officials. Second, members of the implementation team may manipulate

---

[8] Of course, there may also be collusion or kickbacks at the national or district level of the program. This paper, however, focuses on corruption where the bulk of the program money is actually spent—at the village level.

wage payments. As discussed above, villagers in Indonesia typically contribute unpaid or reduced-wage labor to public works projects; corrupt officials can bill the project for labor anyway and pocket the difference. In other cases, those running the project can simply inflate the number of workers paid by the project. All of these types of corruption will be investigated in the empirical work below.

## 3. Experimental Design

The experiments discussed in this paper examine different ways of altering the probability that corruption is detected and punishments are enforced. Three interventions are examined – increasing the probability of external audits ("Audits"), increasing participation in accountability meetings ("Invitations"), and providing an anonymous comment form to villages ("Invitations + Comments"). Section 3.1 discusses the overall experimental design. Section 3.2 then discusses the Audit interventions, and Section 3.3 describes the Invitations and Comment interventions. Section 3.4 discusses the timing of the interventions and data collection.

### 3.1. Experimental Design

Table 1 displays the basic experimental design. As shown in Table 1, randomization into the Invitations and Comment Form treatments was independent of randomization into the Audit treatment. In both cases, the treatments were announced to villages after the project design and allocations to each village had been finalized, but before construction or procurement of materials began.[9] Thus, the choice of what type of project to build, as well as the project's design and planned budget, should all be viewed as exogenous with respect to the experiments.

---

[9] In all villages (including control villages), at the village meeting immediately after the final allocations were announced but before construction began, the study enumerator made a short (less than 5 minute) presentation, introducing him or herself and explaining that there would be a study in the village, that each village and project official would be interviewed for data collection, and that the enumerator would be present to record what happened at each of the accountability meetings. In villages receiving a treatment, the only difference was that this introduction was followed by a description of the treatment(s) in that village. The final engineering survey was not

The randomization of Audits and Invitations/Comment treatments differed in several

ways. First, there was a concern that the audit treatment might be likely to spillover from one

village to another, as officials in other villages might worry that when the auditors came to the

subdistrict, their villages might be audited as well.[10] On the other hand, the participation

treatments were much less likely to have similar spillover effects, as the treatment was directly

observable in the different villages early on. Therefore, the randomization for Audits was

clustered by subdistrict (i.e., either all study villages in a subdistrict received Audits or none did),

while the randomization for Invitations and Comment Forms was done village-by-village. The

calculations of the standard errors below are adjusted to take into account the potential

correlation of outcomes in villages within a subdistrict.

This difference in clustering also necessitated a difference in stratification. As the

Invitations and Comment Forms were randomized village-by-village, they were stratified by

subdistrict, the lowest administrative level above the village. Since the Audits were randomized

by subdistrict, they needed to be stratified at a higher level. Therefore, the Audits were stratified

by district and by the number of years the subdistrict had participated in the KDP program. This

yielded a total of 156 strata for the Invitations / Comment Forms, each containing an average of

3.8 study villages, and 50 strata for the Audits, each containing an average of 3.1 study

subdistricts and 12.1 study villages.

In the analysis, I report three specifications – no fixed effects, fixed effects for each

engineering team that conducted the survey, and stratum fixed effects. Despite the stratification,

the randomization was designed so that the probability each village received a given treatment

---

mentioned at all to the villagers during this presentation, or subsequently, until the surveyors actually appeared to
conduct the survey.
[10] This was most likely to be a problem within subdistricts, as there is frequent communication between both village
officials and project officials within a subdistrict. Communication across subdistrict lines is much more limited,
particularly for village officials.

was always held constant, regardless of what stratum the village was in. The probability of receiving a given treatment is therefore orthogonal to any stratum or village-level variable, so including stratum fixed effects, while it may reduce standard errors, is not necessary for the analysis to be consistent.

Although locations of treatments were randomized by computer according to the procedures described above, it is useful to examine whether, ex-post, they are correlated with village characteristics of interest. I therefore examined the relationship between the treatment variables and ten village characteristics that might be correlated with the outcome variables of interest.[11] As expected given the randomization, these variables are not jointly significant predictors of the treatments. Moreover, including these variables as controls does not affect the results of the experiment presented below.

### 3.2. *The Audit Experiment*
In the "Audit" treatment, villages were told, at the village meeting where they began planning for actual construction, that their project would be audited by BPKP, the government audit agency, with probability 1. They were told that the audit could take place either during or after construction was finished, and would include inspections of both the project's financial records as well as a field-inspection of the construction activities. Approximately two months later, the village implementation team received a one-page letter from BPKP which confirmed that the village had been chosen to be audited and which spelled out in somewhat greater detail exactly what would be covered by the audit.

---

[11] The variables examined are the variables used in analysis of pilot data, and were specified before any of the data used here was collected. The variables include the following: village population, percent of households in the village classified as poor, number of mosques per thousand people, the total budget for the project, the number of sub-projects included, distance from the village to the subdistrict, the village head's age, education, and salary (in hectares of land), and whether the community is mountainous. While several of these variables (village head age, salary, and whether the community is mountainous) are correlated with the audit treatment at the 10% level, the 10 variables considered are not jointly significant.

Villages were told that results of the audits, in addition to being reported to the central government and project officials, would also be delivered directly by the auditors to a special village meeting. Village officials therefore faced several potential sanctions from the audits – retribution from the village, the possibility that the village would not receive KDP projects in the future, and the theoretical possibility of criminal action.

Between one and four months after construction had started, Phase I of the audits commenced.[12] The main purpose of this first round of audits was to credibly demonstrate that the audits were real, rather than an idle threat. One village in each subdistrict receiving the Audit treatment was randomly selected to be audited during this first phase. The audit was conducted over two days, and the results were subsequently presented by the auditors to a specially called public village meeting, where members of the implementation team and village officials were given an opportunity to propose corrective actions for the auditors' findings.

After the first round of audits, all study villages receiving the audit treatment, including the village that was audited during Phase I, were informed in another letter from BPKP that they would be audited again after the construction on the project had been completed. The second phase of the audits was conducted approximately seven months subsequently, after both construction was finished and the collection of the corruption data described below was completed. As with Phase I audits, the results from the Phase II audits were presented to the village in an open village meeting and forwarded to the project for followup.

---

[12] All audits in this phase took place during a three week period during mid January / early February 2004. However, since there was heterogeneity in the timing of when construction started, this was anywhere between 1 and 4 months after construction had begun.

*3.3.   The Participation Experiments*

In the "Invitations" treatment, either 300 or 500 invitations were distributed throughout the village several days prior to each of the three accountability meetings.[13] Though village meetings are officially open to the public, in practice Javanese villagers consider it quite rude to attend a meeting to which they have not been formally invited (usually in writing), and with the exception of a few independent-minded members of the village elite, they rarely do. The village head, who normally issues written invitations for the meetings, therefore has the potential to stack the attendance of the accountability meeting in his favor by issuing invitations only to his supporters. By distributing a large number of invitations, the village head's ability to control who attends the meeting was substantially reduced.

Given the size of a typical village, approximately one in every two households in treatment villages received an invitation. The invitations were distributed either by sending them home with school children, or by asking the heads of hamlets and neighborhood associations to distribute them throughout their areas of the village. The number of invitations (300 or 500) and the method of distributing them (schools or neighborhood associations) was randomized by village.[14]

In the "Invitations + Comment Forms" treatment, invitations were distributed exactly as in the Invitations treatment, but attached to the invitation was a comment form asking villagers' opinions of the project. The idea behind the comment form was that villagers might be afraid of retaliation from village elites, and thus providing an anonymous comment form would increase detection of corruption. The form asked the recipient to answer several questions about the road

---

[13] In addition, for each meeting a small subsidy – Rp. 45,000 ($5) for villages with 300 invitations, Rp. 75,000 ($8) for villages with 500 invitations – was given to the implementation team to cover the additional cost of providing snacks to the extra attendees induced by the invitations.

[14] The purpose of these extra randomizations – the number of invitations and how they were distributed – was to generate additional variation in the number and composition of meeting attendees, to distinguish size effects from composition effects. However, these sub-treatments did not generate a strong enough first-stage to be used in the analysis.

project, and then to return the form – either filled out or blank – to a sealed drop box, placed either at a village school or at a store in the sub-village. The instructions stated clearly that the recipients should not write their name on the form, in order to preserve their anonymity. The form had four closed-response questions (i.e., requesting answers of the form Good, Satisfactory, or Poor) about various aspects of the project, and two free-response questions, one about the job performance of the implementation team and one about any other project-related issues. The comment forms were collected from the drop boxes two days before each meeting and summarized by a project enumerator. The enumerator then read the summary, including a representative sample of the open-response questions, at the village meeting.

## 3.4. Timing

The experiment began in September 2003. After the inter-village forum described in Section 2 made the final allocations of funds, the enumerator went to the village planning meeting that immediately followed and, at that planning meeting, announced any interventions (audits or participation) that would take place in that village. Construction began shortly thereafter, between October and November 2003. Those villages receiving audits received the first letter from BPKP in November 2003, and the first round of audits took place in one randomly selected village in each subdistrict in January 2004, while construction was in progress. The second letter from BPKP was sent out to villages shortly thereafter. The accountability meetings at which the participation interventions were conducted took place after 40%, 80%, and 100% of the funds were spent, between October 2003 and May 2004. The engineering survey to measure missing expenditures took place after construction was finished, between May and August 2004. The final round of audits was conducted in all villages in the audit treatment in September 2004 after all of the data used in this paper had been completed.

## 4. Data

The data used in this paper come from four types of surveys, each designed by the author and conducted specifically as part of the project: a key-informant survey, covering baseline characteristics about the village and the village implementation team; a meeting survey, containing data on the attendees and a first-hand report of discussions at the accountability meetings; a household survey, containing data on household participation in and perceptions of the project; and a final field survey, used to measure corruption in the project. This measurement was conducted in all villages (both treatment and control), and is completely separate from the audits conducted by BPKP as part of the "Audit" treatment. This section describes the final field survey used to measure unaccounted-for expenditures in the road projects; the remaining data, as well as additional details on the field survey, are discussed in more detail in Appendix A.

### 4.1. Reported Expenditures

The key dependent variable I examine is the difference between what villages claim they spent on the project and an independent estimate of what villages actually spent. Obtaining data on what villages claim they spent is relatively straightforward. At the end of the project, all village implementation teams are required to file an accountability report with the project subdistrict office, in which they report the prices, quantities, and total expenditure on each type of material and each type of labor (skilled, unskilled, and foreman) used in the project. The total amount reported must match the total amount allocated to the village. In addition, they also report, for each type of material and labor, the amount donated to the project by villagers. These financial reports were readily available to the survey team for all study villages.

Obtaining an independent estimate of what was actually spent is substantially more difficult, and involves three main activities—an engineering survey to determine quantities of

materials used, a worker survey to determine wages paid by the project, and a supplier survey to determine prices for materials.

## 4.2. *Measuring Quantities of Materials*

In the engineering survey, an engineer and an assistant conducted a detailed physical assessment of all physical infrastructure built by the project in order to obtain an estimate of the quantity of materials used. In the standard road design, known as a *Telford* road, the road consists of three types of materials—a base of sand, a layer of large (10-15 cm) rocks, and a top layer of gravel to provide a smooth running surface.[15] To estimate the quantity of each of these materials, the engineers dug ten 40cm × 40cm core samples at randomly selected locations on the road. By combining the measurements of the volume of each material per square meter of road with measurements of the total length and average width of the road, I can estimate the total quantity of materials used in the road.

It is important to note, however, that this estimate of the materials used in the road, while it should be proportional to the total quantity of materials used in the road, may be smaller in magnitude than the actual amount of materials used in the road, as some amount of loss is normal during construction and measurement. For example, some amount of sand may blow away off the top of a truck, or may not be totally scooped out of the hole dug by the engineers conducting the core sample. I denote the average percentage of materials lost due to normal construction processes and measurement error but with no corruption as the *loss ratio*.

To deal with these loss ratios, whenever possible I express the measured quantities in log form, so that the average loss ratio will be captured by the constant term and will not affect

---

[15] Three other similar road designs are also included in the study. *Telasah* roads are similar to Telford, but install the rocks flatside-up to create a smooth running surface, and therefore largely omit the gravel layer. *Sirtu* roads consist of gravel only, omitting the sand and rock layer. *Katel* roads are similar in design to Telford, but use a mixture of clay and gravel in the top layer to create a more permanent top surface. Telford roads, however, account for 86 percent of the road projects in the sample. The type of road is chosen before the randomization is announced.

estimated differences across villages. However, in some cases, such as when combining different types of materials into the aggregate percent missing measure, this approach is not sufficient, and one actually needs to estimate these loss ratios. One also needs an estimate of these loss ratios if one is interested in the level of the percent missing variable, not just the differences across villages.

To obtain such an estimate, I constructed four short (60m) "test roads" in different areas of East and Central Java as a calibration exercise. During the construction of each of these roads, the survey team carefully measured all quantities before construction (i.e., while still in the delivery trucks). After construction was completed, the techniques described above were used several times, by different engineers, to estimate the quantity of materials used in the road. To allow time for materials to settle and to account for the effects of weather, these followup measurements were conducted anywhere from 1 week to 1 year after the test road was completed. The ratio between the amount of materials actually used in the road and the amount measured after the road was built is an estimate of the loss ratio. I describe this calibration exercise and the resulting loss ratios in more detail in Appendix B.

While the road project comprises the main use of KDP funds in each village, roads are often accompanied by smaller ancillary projects, such as culverts, retaining walls, and gabions, and occasionally by larger projects, such as a small bridge. For each of the ancillary projects, the engineer on the survey team conducted a detailed field survey, measuring and sketching each constructed piece of infrastructure to estimate the volume of materials, such as cement, rocks, and sand, used in the construction.

## 4.3. Measuring Wages and Hours Worked

Workers, defined as people who worked on the project for pay, were asked which of the many activities involved in building the road were done with paid labor, voluntary labor, or some

combination, what the daily wage and number of hours worked was, and to describe any piece

rate arrangements that may have been part of the building of the project. To estimate the quantity

of person-days actually paid out by the project, I combine information from the worker survey

about the percentage of each task done with paid labor, information from the engineering survey

about the quantity of each task, and assumptions of worker capacity derived both from the

experience of field engineers and the experience from building the test roads. These assumptions

of worker capacity are discussed in more detail in Appendix B.

*4.4. Measuring Prices*
       Since there is substantial variation in the prices of construction materials across

subdistricts, a price survey was conducted in each subdistrict. Since there can be substantial

differences in transportation costs within a subdistrict, surveyors obtained prices for each

material that included transportation costs to each survey village. The price survey included

several types of suppliers—supply contractors, construction supply stores, truck drivers (who

typically transport the materials used in the project), and workers at quarries—as well recent

buyers of material (primarily workers at construction sites).[16] For each type of material used by

the project, between three and five independent prices were obtained; I use the median price

from the survey for the analysis.

*4.5. Measures of missing expenditures*
       From the data on the financial reports, I can calculate the expenditures from the financial

reports, which I hereafter will refer to as *reported* amount, and the expenditures estimated from

the field survey, which I will refer to as the *actual* amount. I define the *percent missing* to be the

---

[16] Furthermore, to reduce the potential for bias induced by surveying the actual suppliers for the project, who may be in collusion with project officials, only survey responses from sources outside a given village are used to construct the prices for that village, and no mention of KDP was made until the end of the interview. It turns out that 27% of those interviewed for the price survey had actually been suppliers to the KDP program; dropping them from the analysis, however, does not affect the results. In fact, restricting the price data to only prices obtained from buyers of materials (i.e., dropping *all* suppliers of materials from the price survey) also does not affect the results.

difference between the log of the reported amount and the log of the actual amount. This variable—the percent missing—is the main measure used in the subsequent analysis.

I use several different versions of the percent missing measure in the empirical analysis. First, I report the percent missing for the four major items—sand, rocks, gravel, and unskilled labor—used in the road project. As shown in Table 2, expenditures on these four items account for 90% of reported expenditures on the road project. As these are the four major sources of expenditure, substantial effort was put into ensuring that these four items were measured as accurately as possible in the engineering survey.

As shown in Table 2, the road project accounts for 77% of total funds spent; a further 15% of funds are spent on ancillary projects that go along with the road, such as culverts and retaining walls. Each of these projects was inspected by the field engineers, generating an estimate of the amount of sand, rock, cement, and labor used in each. The second measure of missing expenditures, major items in the main road + ancillary projects, adds in these expenditures as well.[17] Finally, I report the percent missing separately for materials and unskilled labor in the road project.

## 5. Experimental Results

### 5.1. Estimating Equation
Given the randomized nature of the experiments, estimating their effects is straightforward. I estimate an equation of the following form via OLS:

$$\text{PERCENTMISSING}_{ijk} = \alpha_1 + \alpha_2 \text{ AUDIT}_{jk} + \alpha_3 \text{ INVITATIONS}_{ijk} +$$

$$\alpha_4 \text{ INVITATIONSANDCOMMENTS}_{ijk} + \varepsilon_{ijk} \qquad\qquad (1)$$

---

[17] The number of observations is higher when these ancillary expenditures are included because, for some villages, reported expenditures for the main road were combined with those from the ancillary projects. For those villages, the corruption measure could not be constructed for just the main road, but it could be computed for the main road + ancillary projects.

where $i$ represents a village, $j$ represents a subdistrict, and $k$ represents a stratum for the audits. Since the AUDIT treatment variable is perfectly correlated within subdistricts, the standard errors are adjusted to allow for correlation within subdistricts. As each of the 12 engineering teams may have conducted the corruption measurements slightly differently, I estimate a version of equation (1) that includes engineering team fixed effects. Finally, when investigating the audits, I estimate a version of equation (1) that includes fixed effects for each audit stratum $k$, and when investigating the invitations and comment forms, I estimate a version of equation (1) that includes fixed-effects for each subdistrict $j$ (i.e., the stratifying variable for the participation experiments).[18]

## 5.2. *The Audit Experiment*

### 5.2.1. *Overall Effects*

Table 3 presents the main results from the audit experiment. Each row presents the percent missing in different aspects of the project. The first column presents the mean percent missing in the control villages – i.e., those villages that did not receive the audits – and the second column presents the mean level in the villages that received the audits. The effect of the audits – i.e., the coefficient $\alpha_2$ in equation (1) – is presented in column (3). The p-value from a test that the audit effect is zero is presented in column (4). Columns (5) and (6) again present the audit effect and p-values allowing for engineer fixed effects, and columns (7) and (8) present the results allowing for stratum fixed effects.

---

[18] Note that approximately 12% of the observations in the sample were dropped, because the reported expenditure could not be accurately matched to the data from the engineering survey. This was caused by one of four reasons: (1) surveyor error in locating the road, (2) the project consisted largely of a partial rehabilitation of an existing road, (3) agglomerated expenditures reports (i.e., the village expenditure report combined expenditures in the road project with other projects that could not be independently measured, such as a school), or (4) villages that had asphalted the road that refused to let the engineers break the asphalt to conduct the engineering survey. The probability of an observation being dropped for any one of these reasons appears empirically to be unrelated to any of the treatments.

The results show that the audits had a substantial, and statistically significant, negative effect on the percentage of expenditures that could not be accounted for. Column (3) shows that the audits reduced the percent missing in the road project by 8.5 percentage points, and the percent missing in the road and ancillary projects by 9.1 percentage points. These effects are statistically significant, with p-values of 0.058 and 0.034, respectively. Looking across the other columns, these effects are generally of similar magnitude and statistical significance with engineer or stratum fixed effects. The only exception is when stratum fixed effects are included on the road expenditures variable (Coulmn 7, row 1), in which case the estimated effect of the audits is only a 4.8 percentage point reduction.[19] Looking across all of the specifications shown, I conclude that the audits reduced missing expenditures by an average of about 8 percentage points. Compared with the baseline level of about 28 percentage points in control villages, the point estimates imply a reduction in missing expenditures of about 30 percent of the baseline level, although, as discussed above, the absolute levels of the percent-missing variable depend on assumptions for loss ratios and should be interpreted with caution.

Breaking down the change in corruption into materials and labor, Table 3 shows substantial reductions in both materials (sand, rocks and gravel) and unskilled labor associated with the audits, though these separate effects are not statistically significant.[20] Interestingly, in results not reported in the table, I find no significant differences in the effect of the audits

---

[19] The reason that the results are different when stratum fixed effects are included is that doing so effectively removes 13 strata from the sample. The audit randomization was conducted before the list of villages with road projects was known (though the randomization results were kept strictly secret). Out of the initial 166 subdistricts included in the randomization, only 156 subdistricts ended up having villages with road projects. This, plus dropping observations for the reasons described in footnotes 17 and 18, led to 13 out of 50 strata with either all audit or all non-audit villages that are effectively dropped by stratum fixed effects. Estimating the overall results dropping these 13 strata, but without any fixed effects, yields results similar to stratum fixed effects results. This problem is less severe for the major items in roads and ancillary projects, in part because fewer observations are missing and thus fewer strata are effectively dropped.

[20] It is worth noting that, due to the log transformation, mechanically the change in log (materials + wages) and log(materials) + log(wages) will not be identical.

21

between those villages audited both during and after construction and those villages audited only after construction was finished (and, therefore, after the engineering survey was completed). This suggests that the reduction in missing expenditures was caused by the threat of an audit, rather than corrective actions imposed by the auditors once they arrived.

Mechanically, unaccounted-for funds must be accounted for by differences in either the price charged per unit or in differences in the quantities used. When the goods being procured are commodities (as they are in this case), it is much easier for monitors to verify the unit price than to verify the quantity of materials used, so one might expect corruption to occur by inflating quantities rather than prices. To investigate this, in Table 4, I decompose the results into differences in prices and differences in quantities. Column (1) shows that even in control villages, there is almost no difference between reported and actual prices, and that all the unaccounted-for expenditures were due to differences between reported and actual quantities.[21] Consistent with there being no mark-ups of prices to begin with, Columns (3) – (8) show that all of the reductions in missing expenditures caused by the audits were on the quantity dimension.

An important question is whether the observed effects of the audits actually represent a reduction in corruption per se, or whether incompetent builders are simply being replaced by more skilled builders in response to the audits. To investigate this, I examined a number of quality measures, such as compactness of the road, the size and shape of the rocks, and the grade of the road, all of which are relatively inexpensive. Overall competence at construction should affect both expensive and inexpensive components, whereas a reduction in corruption would affect only the expensive aspects of construction quality (i.e. the volume of materials). In fact, I find that controlling for these inexpensive quality measures, either individually or aggregated

---

[21] This difference is statistically significant – a t-test in control villages rejects the equality of differences in prices and differences in quantities with a p-value of less than 0.01.

into an index, does not change the corruption results presented above.[22] This suggests that the results are actually being driven a reduction in corruption per se, rather than an overall change in the competence of those building the project.

### 5.2.2. *Interactions with Village Politics*

An important question is the degree to which the monitoring probability and the punishment level are complements or substitutes.[23] In this context, one important part of the threat posed by an audit is that the village head might lose his re-election bid if corruption was exposed.[24] This may be particularly true for those village heads for whom the previous election for village head was close, and therefore for whom a small change in the vote would be more likely to result in defeat. One can therefore examine whether village heads facing a tough upcoming re-election battle respond more or less to audits, as a way of examining the interaction between the monitoring probability and the expected punishment.

To examine this, I define a village as having an upcoming election for village head if the next such election is scheduled within two years of time of the first audits. I define a narrow victory in the last election as the difference between the runner-up's vote share and the victor's vote share in the previous village head election, so that higher (i.e. less negative) numbers indicate a closer election. I then re-estimate equation (1), including these variables and the interaction of these variables with the audit treatment, where the dependent variable is the percent missing in the road project.

---

[22] Alternatively, putting an index of these variables on the left-hand side of equation (1) shows that there is no change in these inexpensive quality measures in response to the audits.
[23] It is worth noting that the sign of the second derivative of corruption with respect to the auditing probability and the punishment is theoretically ambiguous. See Di Tella and Shargrodsky (2003) for a model that makes this point.
[24] As discussed above, criminal corruption proceedings are becoming more common in Indonesia, but prosecutions against village-level officials, however, remain rare. (Woodhouse et. al, 2004) Electoral and social sanctions against village-levels are therefore likely to be a substantial portion of the threat created by audits.

The results are presented in Table 5. The results suggest that the audits have a substantially larger effect – reducing the percent missing by 31 percentage points – in villages with an upcoming election. When I include the triple-interaction of audits, an upcoming election, and a narrow victory, I find that the effect of the audits is most pronounced for village heads who are up for re-election and whose last election was close – i.e., the village heads who presumably would find face the largest effective punishment for corruption being detected. In results not reported, I find that the effect disappears if the election is more than 2 years away, which suggests that the electorate is relatively myopic. Together, these results suggest that the audits had a greater effect when the punishment from being detected was greater, and that auditing and punishments – and specifically, punishments through democratic accountability – may be complements.

### 5.2.3. *Employment of family members*

The BPKP auditors examined the project's financial records and inspected the construction site. They did not, however, examine who worked on the project, and whether those who worked had family ties to the officials running the project. To the extent that giving jobs to family members is an alternative, but less desirable, method of extracting rents from the project, we might expect this type of non-audited behavior to respond to the increased audits. Of course, there are other reasons to hire family members besides rent-extraction; for example, work on the project by family members might increase if project officials facing audits wanted to improve the project and if family members are less prone to moral hazard or have higher skill than non family members.

I examine the change in employment by family members using data from the household survey described in more detail in Appendix A. Each respondent in the household survey was

asked if he or she was related to any of seven types of village government members or the head of the project implementation team. Overall, approximately 30 percent of respondents are related to some member of the village government, and 6 percent of respondents are related to the head of the project. I examine whether people who said that they were either immediate or extended family members of village government were more or less likely to report having worked for pay on the road project in audited villages than in control villages. Specifically, I estimate the following linear probability model using OLS:[25]

$$\text{WORKED}_{hijk} = \gamma_k + \gamma_2\, \text{AUDIT}_{jk} + \gamma_3\, \text{FAMILY}_{hijk} + \gamma_4\, \text{AUDIT} \times \text{FAMILY}_{ijk}$$

$$+ \gamma_5\, X_{hijk} + \varepsilon_{hijk} \qquad\qquad (2)$$

where FAMILY is a dummy equal to 1 if the individual was a family member of village government or the head of the project, WORKED is a dummy equal to 1 if a household member worked for pay on the project, X is a vector of control variables (age and gender of respondent, predicted household income, dummies for the ways the household was sampled, and the number of social activities household members participated in during the previous three months), $h$ represents the household, $i$ the village, $j$ the subdistrict, $k$ represents the audit stratum, and $\gamma_k$ is a stratum fixed effect. The coefficient of interest is $\gamma_4$, which represents the differential probability in audited villages relative to control villages that family members of village government or the head of the project worked on the project. The empirical results include two different FAMILY variables, one for being a family member of a government official and one for being a family member of the project head.

The results are reported in Table 6. Column (1) shows that family members of government members are 8 percentage points more likely to work on the project in audited

---

[25] Regressions using a Probit specification produce essentially similar results.

villages than in control villages, and column (2) shows that family members of the project head are 14 percentage points more likely to work on the project in audited villages than in non-audited villages. Given that the mean probability of working on the project is only 30 percent, these effects are quite large in magnitude.[26]

As discussed above, there are different interpretations for these results – one view says that this is alternative, less desirable form of corruption, whereas the other suggests this is actually an attempt to improve the project. Though distinguishing between these alternative hypotheses is difficult, there is some suggestive evidence in favor of the nepotism-as-corruption view. In particular, the micro-finance literature has suggested that social connections can be an effective mechanism for minimizing moral hazard (Karlan 2004), so if reducing moral hazard was the issue, one might expect similar effects for workers with many social connections as for family members. In column (3) of Table 6, however, I find that while workers with many social connections are more likely to work on the project overall, there is no statistically significant differential effect in response to the audits in the relationship between social connections and working on the project. Column (4) shows that family member results are still present when I examine all the interactions jointly. Furthermore, in results not reported here, I find that, conditional on observables, family members of village officials are more likely to be employed in the higher wage category (skilled labor rather than unskilled), suggesting that they may be receiving rents from the project. While this evidence is suggestive of a nepotism-as-corruption story, it is by no means definitive, and understanding this phenomenon is an important direction for future work.

---

[26] In results not reported, I find that there were no statistically significant changes in family members working on the project associated with the invitations treatments. For the comment forms treatment, there was a statistically significant increase in family members working on the project, but only for family members of the project head.

*5.3.   The Participation Experiments*

*5.3.1.   Did the interventions increase participation?*

Before we can assess the impact of increased participation on corruption, it is important to make sure that the treatments —i.e., the invitations and comment forms – did, in fact, increase villager participation. To examine this, in Table 7 I re-estimate equation (1), where the dependent variables are several measures of participation in the accountability process. Each observation represents one accountability meeting, so there are three observations for each village. Standard errors are adjusted to take into account this intra-village correlation.

Column (1) shows that the treatments had a substantial effect on total attendance – the invitations treatment increased attendance at the meetings by an average of 14.8 people, or approximately 40 percent.[27] The slightly smaller increase for villages receiving comment forms as well as invitations suggests that being able to submit written comments and attending meetings are substitutes. Column (2) shows that virtually all of the increase in attendance at the meetings came in the form of increased attendance by these non-elite villagers, so that the number of non-elite at the meetings increased by 75 percent, from 16 people in the control villages to 26 people in the treatment villages.[28]

Columns (3) and (4) examine the impact of the treatments on active participation at the meetings. Column (3) shows that, in the invitations treatment, the average number of people who spoke at a meeting increased by 0.74, or just under 10 percent. Column (4) shows that about 40 percent of these new speakers were non-elite, increasing the number of non-elite villagers who spoke at a meeting by about 30 percent over the baseline level.

---

[27] Villages receiving 500 invitations had slightly higher attendance (by 1.7 people on average) than villages receiving 300, though these differences are not statistically significant. Passing out invitations through schools did not result in a statistically significant difference in composition of attendees. I therefore omit these additional sub-variants in the subsequent analysis, though including them does not substantially alter the results.
[28] I classify people as "non-elite" if they have no official position in the village, no official position on the project, and were not described as a *tokoh masyarakat* ("informal village leader") by village members who assisted the enumerator.

In addition, the comment forms appear to have been quite successful in eliciting villagers' opinions about the project. On average, 140 comment forms (about 35 percent) were returned, filled out, per meeting. The responses were quite varied, and on average had slightly more positive than negative comments. There are no substantial differences in the response rate across villages of differing average education levels—mean adult education in these villages is 4.8 years, and 89 percent of adults can read and write—which suggests that, at least within the level of education in rural Java, education does not seem to be a substantial constraint to using comment forms to elicit villager responses.

### 5.3.2. *Effect on Meetings*

Table 8 investigates the effect that increased participation (via the invitations and comment forms) had on the accountability meetings. As discussed in more detail in Appendix A, the enumerator recorded each problem or issue that was discussed at the meeting, and coded whether the problem was potentially corruption-related or not.[29] In addition, I define a 'serious response' to a problem as any of the serious actions that could be taken by a village in response to a problem with the project – specifically, agreeing to replace a supplier or village official, agreeing that money should be returned, agreeing for an internal village investigation, asking for help from district project officials, or requesting an external audit. These serious actions are quite rare – they occur at only 3 percent of meetings – and thus, to preserve statistical power, I consider them together.

The results in Column (1) suggest that neither the invitations treatment nor the invitations + comment form treatment had a significant effect on the total number of problems discussed at

---

[29] Classifying problems and listing whether they are potentially corruption related clearly requires some degree of subjective judgment on the part of the enumerator filling out the form. However, all villages in a subdistrict were handled by the same enumerator, so including stratum (i.e., subdistrict) fixed effects controls for these potential differences in coding.

the meeting. This implies that the increase in the number of people talking in Table 7 is an increase in who participates in the discussion, not an increase in the number of problems per se. However, as shown in Column (2), both the invitations and the invitations + comment forms increased the probability of having a corruption-related problem discussed at the meeting by 2.7 percentage points, or 50 percent above the baseline level. Only the comment forms, however, affected how problems were resolved – column (3) shows that the probability of a serious action being taken is 1.5 percentage points higher – or 70 percent higher than the baseline amount – in villages receiving the comment forms, but that there is no effect in villages receiving only invitations.

These results suggest that the impact of the comment forms was slightly different from what was expected. In particular, despite the large number of comments received, adding the comment forms did not change the probability that a corruption-related problem was discussed. This does not mean that fear of retaliation was not an issue, however. Rather, the results suggest that when an issue was brought up, possibly by an elite member of the village who had less fear of retaliation by the implementation team, villagers may have been unsure whether to side with the challenger or the implementation team. Knowing from the comment forms that many other people agreed with them, and therefore that the challenge was likely to be victorious, may have tipped the balance. This suggests the comment forms were more about creating common knowledge about problem, rather than bringing previously unknown problems to light. Nevertheless, both of these effects were small in absolute magnitude.

### 5.3.3. *Effect on Missing Expenditures*

Table 9 examines the overall impact of the two participation treatments on the percent missing in the projects. The first panel shows the effect of the invitations treatment; the second

panel shows the effect of the invitations + comment forms treatment. The results suggest that both the invitations and invitations + comment forms treatments had a small, and statistically insignificant, impact on the overall percent missing from the project. Depending on the specification and the measure of corruption, the point estimates suggest that these treatments reduced the percent missing by between 1.5 and 3 percentage points, though these estimates are never statistically distinguishable from 0.

The lack of a strong effect of the invitations and comment forms on corruption is consistent with the evidence in Section 5.3.2, for while the comment forms did increase the probability of serious action being taken in response to corruption, the magnitude of this response was small. In fact, even if taking a serious anti-corruption action at a meeting eliminated corruption entirely, the reduction in corruption in caused by these induced serious actions would have been less than 1 percentage point of average total expenditures.[30] The results here suggest that not only were the direct effects (via anti-corruption actions) small, but any deterrent effects of the treatments were small as well.

However, the small effects on the overall percent missing variable mask the fact that there were substantial reductions in missing labor expenditures. As shown in Table 9, the invitations treatment led to a statistically significant reduction in missing expenditures in labor of between 14 percentage and 18 percentage points. The invitations + comment forms also led to a reduction in missing labor expenditures, though it was somewhat smaller in magnitude (9 to 13

---

[30] To see this, suppose that taking a serious anti-corruption action eliminates corruption in the village. (In fact, in the cross-section, villages where a serious response was taken had a missing expenditures about 18 percentage points lower than other villages.) If I re-estimate Table 8 at the village level, rather than the meeting level, the point-estimate is that comment forms treatment increased the probability of a serious response to a problem by 3.8 percentage points. This suggests that, if the sole effect of the comment forms was through the probability of a serious response being taken, the expected average reduction in the percent missing would be $18 \times 0.038 = 0.68$ percentage points.

percentage points) and not statistically significant.[31] On the other hand, the point estimates suggest no change, or if anything a very mild increase, in missing funds on the materials dimension. Since, as shown in Table 2, materials account for 68% of road expenditures whereas unskilled labor accounts for only 20%, the lack of an effect on the materials dimension is why the effects on the overall percent missing are small and statistically insignificant despite the reduction in missing labor.

There are several different potential explanations for the reduction in labor expenditures but not materials associated with the participation treatments. First, it may be easier for villagers to observe actual wage payments than the quantity of materials delivered, making corruption in labor technologically easier for villagers to detect. Alternatively, even if villagers had equal information about both types of corruption, the focus on labor may have arisen because the invitations induced more workers to attend the meetings, and those workers focused on their private interest (i.e., the wages they personally were supposed to be paid by the project) rather than the public good of a higher quality road.

Though these hypotheses are difficult to definitively distinguish, there is suggestive evidence in favor of the hypothesis that it was more about incentives than information. First, the household survey described in Appendix A included a question on the respondent's perceptions of corruption in the road project. In related work (Olken 2005), I find that these perceptions of corruption in the project were positively correlated with missing materials expenditures, but were uncorrelated with missing labor expenditures. This suggests that, surprisingly, villagers appear to have more information about missing materials expenditures, not missing labor expenditures.

---

[31] I cannot reject the hypothesis that the two participation treatments (invitations and invitations + comment forms) had the same effect. Pooling them together yields an average reduction in labor expenditures of between 12 and 17 percentage points, with the average reduction statistically significant in 2 of 3 specifications.

Second, one can examine whether the invitations treatment varies when workers come from outside the village, and are therefore ineligible to attend the accountability meetings. When I do this (i.e., interact the invitations treatment with the percent of workers coming from outside the village), I find that the reduction in missing labor associated with the invitations treatment is smaller when more of the workers come from outside the village. Since the only people that attend the accountability meetings are those who live in the village, this suggests that it is the fact that the invitations induce workers to attend meetings and fight for their own wages that drives the reduction in theft in wages. The percent of workers coming from outside the village is potentially endogenous, so these results should be viewed as speculative. Nevertheless, this result, combined with the result about information, suggests that the lack of a reduction in materials may be because monitoring theft of materials is more of a public good than monitoring theft of wages.

Even without the invitations and comment forms, the KDP program investigated here includes more grass-roots participation, and a more complicated system of checks and balances, than the typical government project in most developing countries. It is possible that the marginal effects of increasing participation from this relatively high baseline—i.e., the effects reported above—do not capture the overall effects of grass-roots participation as a monitoring mechanism, and that more dramatic variations in the amount of participation might have different effects. However, the fact that the invitations did lead to a substantial reduction in missing labor expenditures suggests that the changes in participation induced by the interventions were large enough to make a difference.

## 6. Cost-Benefit Analysis

The previous analysis discussed the direct effects of the treatments. This section performs a cost-benefit calculation to assess whether, on net, the benefits from these treatments exceed the costs. This exercise requires making several assumptions, particularly about the efficiency cost of different types of missing expenditures, and therefore should be viewed as somewhat more speculative than the preceding sections.

Table 10 presents the cost-benefit estimates for the audits and invitations treatments. (Since none of the estimated effects of invitations + comment forms were statistically significant, I omit them from this analysis.) For each treatment, I present two sets of net benefits – "equally weighted net benefits," calculated under the assumption that the marginal utility of income is constant across individuals, and "distribution weighted net benefits," which takes into account the fact that some benefits and costs are borne by the rich while others are borne by the poor.[32]

The net benefits are derived as follows. First, the monetary cost of the audits is the actual cost paid by the project per audit, including the salary of the auditors, and the monetary cost of the invitations is the actual cost paid for photocopying the invitations. The associated dead-weight loss is the dead-weight loss associated with the increased taxes required to pay the monetary cost of the treatments.[33] Finally, the time cost for both treatments is the monetary value

---

[32] Specifically, the distribution-weighted net benefits assumes CRRA utility of per-capita consumption with a coefficient of relative-risk aversion of 2, normalized so that the median household in rural Java has marginal utility of 1. Households where someone is in the village government have per-capita expenditure 18.5 percent higher than typical households in the village, so this (plus the CRRA assumptions) suggests that the social value of $1 of rents received by project or government officials is $0.61. On the other hand, the social value of $1 of increased wages received by workers is $1.29, as workers have per-capita consumption approximately 13 percent lower than the median in the village. I assume that the social benefits from the road are enjoyed equally by all in the village, so they have a marginal social value of 1. Finally, for taxes, I estimate the distributional impact of both the monetary and deadweight loss cost using national consumption data from the 2003 SUSENAS, assuming that the burden is borne proportionally to consumption.

[33] Estimates of the marginal cost of public funds for indirect taxes vary considerably. For developing countries, estimates for the marginal cost of funds from indirect taxes range from 1.04 to 1.05 in Indonesia and Bangladesh (Devarajan et. al, 1999) to 1.59 to 2.15 for India (Ahmad and Stern, 1987). By comparison, estimates of the marginal cost of public funds for the U.S. range from 1.17 to 1.56 (Ballard et. al, 1985), with policy analysis

of the additional time villagers spend at village meetings as a result of the treatments, valued at average local wage rates.

The estimates for the change in rents received by corrupt officials are taken from Column (3) in Table 3 and Table 9, multiplied by the average cost of the project and the average percent of the project consisting of materials and labor expenditures from Table 2. To estimate the change in benefits from the project, I use the point estimates of the change in actual expenditures relative to the original (pre-randomization) plan.[34] For wages, the increase in actual wage expenditures is valued at the actual amount transferred. For materials, the increase in materials expenditure increases the lifespan of the road. To value the change in lifespan of the road, I use estimates from a cost-benefit evaluation of KDP roads conducted for the World Bank (Dent 2001), which imply that the marginal dollar of materials stolen reduces the discounted benefits from the road by 3.41 dollars.[35] For other reported expenditures, I assume a marginal social value of 1, since it is unclear what percentage of them are actually spent productively on the road as opposed to being captured by local officials.

The results presented in Table 10 suggest that the audits were substantially cost-effective. I estimate that the *net* social benefits from the audits were approximately $250 per village, which

---

typically using values in the range from 1.30 to 1.40. I assume that the marginal cost of public funds is 1.4 (i.e. the dead-weight loss is 0.4)

[34] The reason to use the change in actual expenditures relative to the original plan, rather than missing expenditures in each category, is that there may have been changes in where corruption was hidden in the accounting ledgers, and such nominal changes in reporting would have no efficiency consequences. For example, in results not reported in the table, I find that the reduction in missing labor expenditures was due to a combination of shifting reported expenditures – there is a reduction in reported labor expenditures and a corresponding increase in reported other expenditures (i.e., the 12% of expenditures on things other than rocks, sand, gravel, and unskilled labor) – as well as a real increase in actual wage payments to workers, though these changes are not statistically significant. The changes in missing expenditures due to the audits, however, all show up as increases in actual expenditures.

[35] Dent estimates that KDP road projects produces annual flow benefits of 33% of their cost each year for the life of the road. He also estimates that a good quality road will last 10 years and a poor quality road will last 5 years. Based on conversations with KDP engineers, I assume that each percentage of materials stolen reduces the life of the road by 0.1 years, so stealing 50% of materials is the difference between 'good' and 'poor' quality. Assuming a real discount rate of 5% and that the baseline lifespan of a road is 7 years implies that each dollar of materials stolen reduces the discounted net benefits of the project by 3.41 dollars.

implies that the benefits were more than 150% the cost of the audits. Weighting by distributional incidence increases these net benefits even more. For the invitations treatment, I estimate net benefits that are close to zero, since a reduction in theft of labor is just a transfer from wealthy village leaders to ordinary workers.

The audit treatment discussed here was a move from a 4% baseline audit probability to a 100% audit probability. It is possible, however, that the response of corruption to the audit probability is concave, so that an audit probability of only 50% or even 25% might achieve most of the benefits.[36] The costs of audits, however, are roughly linear in the audit probability. This suggests that raising audit probabilities to an intermediate level, rather than all the way to 100%, might have an even higher cost-benefit ratio than the 100% audit probability documented here.

## 7. Conclusion

This paper has examined the results of a field experiment in Indonesia, designed to investigate alternative approaches to fighting corruption. I examined the effect of two strategies: top-down monitoring by government auditors, and bottom-up monitoring through grass-roots participation in the village monitoring process. The evidence suggests that increasing the probability of external audits substantially reduced missing funds in the project. A cost benefit analysis suggests that the benefits from the audits exceeded their cost.

By contrast, increasing grass-roots participation in monitoring the project affected only missing labor expenditures, with no impact on materials and, as a consequence, little impact overall. These results suggest that grass-roots monitoring may be more effective for government programs that provide private goods, such as subsidized food, education or medical care, where individual citizens have a personal stake in ensuring that the goods are delivered and that theft is

---

[36] For example, Nagin. et al. (2002), in a study of monitoring of call center employees, found substantial evidence of diminishing returns to increasing the monitoring probability.

minimized. For public goods where incentives to monitor are much weaker, such as the infrastructure projects studied here, the results suggest that using professional auditors may be much more effective.

This does not mean that empowering community members to discipline service providers has no role in an anti-corruption approach. In fact, the results suggests that the audits were most effective when the village head was up for re-election, which suggests that local level accountability may be an important mechanism for disciplining public officials. However, the results suggest that grass-roots monitoring alone may not be sufficient, and that for detecting corruption professional monitors may also important.

The results in this paper present the results from a short-run intervention. If auditors are bribable, over time villages may develop repeat relationships with auditors which may make bribing auditors easier than in the one-shot case examined here. This might suggest, for example, that frequent rotation of auditors – or lower probabilities of audits combined with higher punishments – may be optimal. Understanding the long-run implications of anti-corruption policies, and developing ways of ensuring that they remain effective over the long term, is an important issue for future research.

# References

Ahmad, Etisham and Nicholas Stern, "Alternative Sources of Government Revenue: Illustrations from India, 1979-80," in David Newbery and Nicholas Stern eds., *The Theory of Taxation for Developing Countries*, New York: Oxford University Press, 1987.

Ballard, Charles L., John B. Shoven, and John Whalley, "General Equilibrium Computations of the Marginal Welfare Cost of Taxes in the United States," *American Economic Review 75 (1)*, pp 128-138, March 1985.

Bardhan, Pranab, "Decentralization of Governance and Development," *Journal of Economic Perspectives 16 (4)*, pp. 185-205, Fall 2002.

Bardhan, Pranab and Dilip Mookherjee, "Corruption and Decentralization of Infrastructure Delivery in Developing Countries," *Economic Journal*, forthcoming..

Becker, Gary S. and George J. Stigler, "Law Enforcement, Malfeasance, and Compensation of Enforcers," *The Journal of Legal Studies 3 (1)*, pp. 1-18, January 1974.

Cadot, Olivier, "Corruption as a Gamble," *Journal of Public Economics 33*, pp. 223-244, 1987.

Dent, Geoffrey, "Ex-Post Evaluation of Kecamatan Development Program (KDP) Infrastructure Projects," mimeo, Project Appraisals Pty Limited, November 2001.

Deaton, Angus, *The Analysis of Household Surveys*, Baltimore: Johns Hopkins University Press, 1997.

Devarajan, Shantayanan, Karen E. Thierfelder, and Sethaput Suthiwart-Narueput, "The Marginal Cost of Public Funds in Developing Countries," in Amedeo Fossati and Wolfgang Wiegard (eds.), *Policy Evaluations with Computable General Equilibrium Models*, New York: Routledge Press, 2002.

Di Tella, Rafael and Ernesto Schargrodsky, "The Role of Wages and Auditing During a Crackdown on Corruption in the City of Buenos Aires," *Journal of Law and Economics 46*, April 2003.

Fisman, Raymond and Shang-Jin Wei, "Tax Rates and Tax Evasion: Evidence from 'Missing Imports' in China," *Journal of Political Economy 112 (2)*, pp. 471-500, 2004.

Hsieh, Chang-Tai and Enrico Moretti, "Did Iraq Cheat the United Nations? Underpricing, Bribes, and the Oil for Food Program," NBER Working Paper #11202, 2005.

Karlan, Dean S., "Social Connections and Group Banking," mimeo: Princeton University, 2004.

Krueger, Anne O., "The Political Economy of the Rent-Seeking Society," *American Economic Review 64 (3)*, pp. 291-303, June 1974.

Lambsdorff, Johann G., "Background Paper to the 2003 Corruption Perceptions Index," mimeo: Transparency International, September 2003.

Mauro, Paulo, "Corruption and Growth," *Quarterly Journal of Economics 110 (3)*, pp. 681-712, August 1995.

Olken, Benjamin A., "Corruption and the Cost of Redistribution," *Journal of Public Economics*, forthcoming.

Olken, Benjamin A., "Measuring Corruption: Perceptions vs. Reality," mimeo, Harvard University, 2005.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor, "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment," *American Economic Review 92 (4)*, pp. 850-873, September 2002.

Reinikka, Ritva and Jakob Svensson, "Local Capture: Evidence from a Central Government Transfer Program in Uganda," *Quarterly Journal of Economics 119 (2)*, pp. 679-706, May 2004.

Rose-Ackerman, Susan, "The Challenge of Poor Governance and Corruption," mimeo, Copenhagen Consensus and Yale University, 2004.

Shleifer, Andrei and Robert W. Vishny, "Corruption," *Quarterly Journal of Economics 108 (3)*, pp. 599-617, August 1993.

Stiglitz, Joseph E., "Participation and Development: Perspectives from the Comprehensive Development Paradigm," *Review of Development Economics 6(2)*, pp. 163-192, 2002.

Woodhouse, Andrea et. al., "Village Justice in Indonesia: Case Studies on Access to Justice, Village Democracy, and Governance," mimeo, World Bank Indonesia Office, February 2004.

*World Development Report 2004: Making Services Work For Poor People*, Washington, DC: World Bank, 2004.

Yang, Dean, "Can Enforcement Backfire? Crime Displacement in the Context of a Common Customs Reform," University of Michigan, mimeo, 2004.

**Appendix A   Data collection**

In addition to the corruption field survey described in Section 4 above, this paper uses three other types of data collected during the course of the project—key-informant surveys, data on village meetings, and a household survey. This Appendix describes the data on village meetings and the household survey. In addition, it provides additional details on the field survey described above.

*Village meeting data*

In each KDP year in a village, there are a total of seven regularly scheduled village meetings that must occur (including the three accountability meetings.) The enumerator was present at the final four meetings – the meeting at which preparations for construction were begun plus the three accountability meetings. At each meeting, the enumerator circulated an attendance list. As the meeting progressed, the enumerator (with the assistance of a local counterpart) noted on the attendance list each person who spoke at the meeting. In addition, sitting with the local counterpart after the meeting was over, the enumerator asked the local counterpart to identify which of the attendees was a *tokoh masyarakat*, or informal leader, a designation typically given to teachers, religious leaders, or other types of informal village leaders.

While the meeting was in progress, the enumerator was asked to keep detailed notes on what occurred during the meeting. The enumerator compiled a list of all problems that arose at the meeting. A "problem" was defined as the topic of any substantial discussion other than the routine business of the meeting; the median problem reported in the data was discussed for 7 minutes, and the mean number of problems reported in an Accountability Meeting was 0.73. For each problem, the enumerator described the problem, classified it according to one of 57 pre-defined problem codes, and listed the amount of time spent discussing the problem, who first raised the problem, who was potentially involved in the problem, whether there were indications of corruption in the problem, whether the problem was resolved, and if so, what actions were taken to resolve it.

*Household survey*

The household survey was conducted approximately during the last two months of construction and the first month after construction was completed. The survey contained a household roster, a list of assets, information on participation in social, religious, and government activities, detailed information on participation in the road project, and a series of questions about perceptions of corruption. Household expenditures were predicted based on assets, using the relationship between assets and consumption from the 1999 SSD (Hundred Villages Survey). The household survey was designed as a stratified random sample, containing between six and thirteen respondents per village.

*Field survey / Corruption measurement*

The general approach used in the field survey is described in Section 4 of the text above. This section discusses a number of additional aspects of the data collection not discussed above that are important for the analysis in this paper.

One important issue is the treatment of voluntary contributions. According to official village reports, these contributions account for an average of 16% of total project costs. Of these voluntary contributions, the bulk (60% according to official village reports) comes in the form of either voluntary labor or tools (typically, village workers bring their own tools to work on the project; the value of these tools is often reported as a voluntary contribution). Anecdotal evidence confirms that voluntary contributions are qualitatively important primarily for labor and tools.

I treat these voluntary contributions as follows. For the three main materials in the road projects—sand, rocks, and gravel—I include both reported project expenditures and reported voluntary contributions when calculating the total amount of reported materials, as this total amount should be comparable to the total amount of each type of material observed in the field. For labor, where overstating voluntary labor is a much more important potential margin of corruption than in materials, a different approach was taken. As discussed in Section 4, for labor, I use only reports and actual estimates of paid labor, and exclude voluntary labor.

For the worker survey described in Section 4 above, a total of five worker interviews were conducted in each village. Of these interviews, two were focus group interviews, consisting of 3 or more workers interviewed together, and three were individual interviews. Two of these interviews (one focus group, one individual) were workers randomly selected form the official list of all workers who had ever been paid by the project; the remainder were recruited more informally, by having the surveyor go to different areas of the village located near the road and asking households who had worked on the project.

The responses on the worker survey are in some cases quite variable. In particular, there is often variation because certain workers only worked on certain aspects of the project, because certain parts of the project were done differently than others, or because some workers were confused about the meaning of the questions. Because these discrepancies can be quite difficult to capture on the survey, surveyors were asked to fill out a form in which they summarized the results of all interviews. The surveyors were instructed that the summary should represent their best understanding of what actually happened in the village, based only on the information in the worker surveys. Though this summary is necessarily more subjective than answers to the worker surveys, experience during the pilots suggests that this method is more accurate than using mechanical averages from the individual worker surveys. Accordingly, the main results use the information from this summary report.

## Appendix B   Assumptions and Calibration

Two main types of assumptions are used in the corruption calculations in this paper—assumptions about loss ratios, defined as the percentage of materials unaccounted for as a result of normal construction losses and measurement error, and assumptions about worker capacity, defined as the amount of each type of task an average worker can accomplish in a day of work. This section describes the assumptions used in the paper in more details, and discusses the calibration exercises through which these assumptions were determined.

To determine the appropriate assumptions, two methods were used. First, similar assumptions are used in the planning process for KDP roads. For example, KDP engineers typically assume normal construction losses of 16 percent for sand and gravel, and 23 percent for rock. Therefore, when purchasing supplies, engineers increase the final volume of materials they need by these amounts to ensure the appropriate quantities at the time of construction. Similarly, when budgeting labor requirements, KDP engineers estimate the quantities of each type of task

in the project, and multiply by standard worker capacity estimates. As a baseline, I therefore obtained, from the KDP chief engineer, the standard assumptions typically used for KDP projects.

There are several problems with these assumptions, however. For the loss ratios for physical materials, the estimates included only the losses during the construction process; they do not include the additional losses inherent in the measurement process. They are therefore likely to be too small for the purposes here. For manual labor, the estimates used were apparently based on a 1970s workfare program, when both worker nourishment and worker motivation were substantially lower than today.

Given this, I conducted 4 calibration exercises. In each calibration exercise, our project built a 60 meter Telford road, similar in construction techniques and standards to KDP road projects. Materials were purchased from local suppliers, and labor was recruited from the villages in which the roads were constructed. The roads were constructed in four very different regions of the study area, which were chosen to represent the different types of conditions typical in the study area. Detailed measurements were taken of all materials delivered to the site, and careful track was kept of each worker's activities during construction.

Once the road was completed, the measurement techniques from the study were applied to the test road. Using these techniques, I estimated the total quantity of material used in each road. By comparing the quantity of material estimated using the study techniques to the actual quantity of material we used in the road (which is known, since it was measured as it arrived at the project site), one can recover the correct loss ratios.

From the logs of worker activity, I was able to construct the actual time required by the workers on these test roads to complete each of the tasks assigned to them. Of course, workers who know that they are being closely watched may work more quickly than normal workers in the field. Nevertheless, the pace of work in the test road was between 50 and 700 percent faster, depending on the task, than the standard assumptions. Additional conversations with field engineers confirmed that the 1970s standards were, in fact, quite low. As one engineer said, the standards were so loose that "if a project using those standards didn't finish with money left over, I was immediately suspicious that there was corruption in that village."

Table 11 lists the main assumptions used in the study, based on the results of the calibration. For the labor estimates, the revised estimates were revised downward 20% from the estimates from the test road, to incorporate breaks taken by workers and to take into account the fact that carefully watched workers may work faster than workers under more normal monitoring conditions.

Interestingly, the high loss ratio for gravel and low loss ratio for sand suggests that some of the gravel is seeping through cracks in the rocks, and is counted by our survey as sand. This suggests that one might better consider "sand + gravel" together (with an implied joint loss ratio of 1.25), rather than separately. Doing so reduces the average percent missing from 23 percent to 21 percent, but otherwise does not alter the results of the paper.

## Table 1: Number of villages in each treatment category

| | Control | Invitations | Invitations + Comment Forms | Total |
|---|---|---|---|---|
| Control | 114 | 105 | 106 | 325 |
| Audit | 93 | 94 | 96 | 283 |
| Total | 207 | 199 | 202 | 608 |

Notes: Tabulations from results of randomization. Each subdistrict faced a 48% chance of being randomized into the Audit treatment. Each village faced a 33% chance of being randomized into the Invitations treatment, and a 33% chance of being randomized into the Invitations + Comment Forms treatment. The randomization into Audits was independent of the randomization into Invitations or Invitations + Comment Forms.

## Table 2: Summary Statistics

| | |
|---|---|
| Total project size (US$) | 8,875 |
| | (4,401) |
| | |
| Share of total reported expenses | |
| Road project | 0.766 |
| | (0.230) |
| Ancillary projects | 0.154 |
| (culverts, retaining walls, etc) | (0.181) |
| Other projects | 0.079 |
| (schools, bridges, irrigation, etc) | (0.166) |
| | |
| Share of reported road expenses | |
| Sand | 0.099 |
| | (0.080) |
| Rocks | 0.484 |
| | (0.143) |
| Gravel | 0.116 |
| | (0.181) |
| Unskilled labor | 0.196 |
| | (0.125) |
| Other | 0.105 |
| | (0.164) |
| | |
| Number observations | 538 |

Notes: Statistics shown are means, with standard deviations in parentheses.
Data for the 538 villages for which percent missing could be calculated.
Exchange rate: Rp. 9,000 = US $1.

**Table 3: Audits – main theft results**

| *Percent missing:* *Log reported value – Log actual value* | Control Mean | Treatment Mean: Audits | No Fixed Effects | | Engineer Fixed Effects | | Stratum Fixed Effects | | Num Obs |
|---|---|---|---|---|---|---|---|---|---|
| | | | Audit Effect | P-Value | Audit Effect | P-Value | Audit Effect | P-Value | |
| Major items in roads | 0.277 (0.033) | 0.192 (0.029) | -0.085 (0.044) | 0.058 | -0.076 (0.037) | 0.041 | -0.048 (0.033) | 0.144 | 477 |
| Major items in roads and ancillary projects | 0.291 (0.030) | 0.199 (0.030) | -0.091 (0.043) | 0.034 | -0.086 (0.038) | 0.023 | -0.090 (0.036) | 0.012 | 538 |
| Breakdown of roads: | | | | | | | | | |
| Materials | 0.240 (0.038) | 0.162 (0.036) | -0.078 (0.053) | 0.143 | -0.063 (0.042) | 0.141 | -0.034 (0.040) | 0.398 | 477 |
| Unskilled labor | 0.312 (0.080) | 0.231 (0.072) | -0.077 (0.108) | 0.477 | -0.090 (0.088) | 0.310 | -0.041 (0.076) | 0.590 | 426 |

Notes: Audit effect, standard errors, and p-values are computed by estimating equation (1), a regression of the dependent variable on a dummy for audit treatment, invitations treatment and invitations + comment forms treatments. Robust standard errors in parentheses, allowing for clustering by subdistrict (to account for clustering of treatment by subdistrict). Each 'audit effect', standard error, and accompanying p-value is from a separate regression. Each row shows a different dependent variable, shown at left. All dependent variables are the log of the value reported by the village less the log of the estimated actual value, which is approximately equal to the percent missing. Villages are included in each row only if there was positive reported expenditures for the dependent variable listed in that row.

**Table 4: Audits – prices vs. quantities**

| *Percent missing:* *Log reported value – Log actual value* | Control Mean | Treatment Mean: Audits | No Fixed Effects | | Engineer Fixed Effects | | Stratum Fixed Effects | | Num Obs |
|---|---|---|---|---|---|---|---|---|---|
| | | | Audit Effect | P-Value | Audit Effect | P-Value | Audit Effect | P-Value | |
| Prices | -0.018 (0.017) | -0.016 (0.021) | 0.002 (0.027) | 0.949 | 0.011 (0.025) | 0.666 | 0.018 (0.024) | 0.458 | 494 |
| Quantities | 0.276 (0.030) | 0.207 (0.028) | -0.069 (0.041) | 0.092 | -0.069 (0.035) | 0.053 | -0.048 (0.032) | 0.141 | 477 |

Notes: See Notes to Table 3. Reported and actual prices are defined as a weighted average across different commodities (rock, sand, gravel, and unskilled labor), where each commodity is weighted by the reported quantity. Reported and actual quantities are also defined as a weighted average, where each commodity is weighted by the reported prices.

## Table 5: Interactions with village politics

|  | (1) | (2) | (3) |
|---|---|---|---|
| Upcoming election | 0.144 | 0.026 | 0.219 |
|  | (0.111) | (0.092) | (0.167) |
| Narrow victory | 0.092* | 0.049 | 0.024 |
|  | (0.052) | (0.064) | (0.062) |
| Audit | -0.060* | -0.039 | -0.008 |
|  | (0.033) | (0.049) | (0.047) |
| Upcoming election * Audit | -0.306* |  | -0.668** |
|  | (0.176) |  | (0.268) |
| Narrow victory * Audit |  | 0.112 | 0.159 |
|  |  | (0.100) | (0.102) |
| Narrow victory * Upcoming election |  |  | 0.187 |
|  |  |  | (0.218) |
| Narrow victory * Upcoming election * Audit |  |  | -0.878** |
|  |  |  | (0.428) |
| Stratum fixed effects | YES | YES | YES |
| Observations | 428 | 428 | 428 |
| R-squared | 0.38 | 0.37 | 0.39 |
| Mean dep. var | 0.24 | 0.24 | 0.24 |

Notes: Results from estimating equation (1), with additional interactions between the audit variable and the village-level characteristics shown in the table. Dependent variable is missing expenditures in the road project. Robust standard errors in parenthesis, adjusted for clustering at the subdistrict level. Upcoming election is a dummy variable for whether the village head election is scheduled within the next 2 years. Narrow victory is defined as the difference between the vote share of the second place candidate and the winning candidate in the most recent village head election, so that larger (i.e., less negative) numbers indicate a closer election.

## Table 6: Nepotism

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Audit | -0.011 | 0.004 | -0.017 | -0.038 |
|  | (0.023) | (0.021) | (0.032) | (0.032) |
| Village Gov't Family Member | -0.020 | 0.016 | 0.016 | -0.014 |
|  | (0.024) | (0.017) | (0.017) | (0.023) |
| Project Head Family Member | 0.051 | -0.015 | 0.051 | -0.004 |
|  | (0.032) | (0.047) | (0.032) | (0.047) |
| Social activities | 0.017*** | 0.017*** | 0.013* | 0.014** |
|  | (0.006) | (0.006) | (0.006) | (0.006) |
| Audit × Village Gov't Family Member | 0.079** |  |  | 0.064* |
|  | (0.034) |  |  | (0.034) |
| Audit × Project Head Family Member |  | 0.138** |  | 0.115* |
|  |  | (0.060) |  | (0.061) |
| Audit × Social activities |  |  | 0.010 | 0.008 |
|  |  |  | (0.008) | (0.008) |
| Stratum Fixed Effects | YES | YES | YES | YES |
| Observations | 3386 | 3386 | 3386 | 3386 |
| R-squared | 0.26 | 0.26 | 0.26 | 0.27 |
| Mean dep. variable | 0.30 | 0.30 | 0.30 | 0.30 |

Notes: Data is from the household survey. Each observation represents one household. Results from estimating equation (2), where the dependent variable is a dummy for whether a household member worked (for pay) on the road project. Estimation is by OLS with stratum fixed effects. Robust standard errors in parentheses, adjusted for clustering at subdistrict level. Social activities refers to the number of social activities adult household members participated in during the last month. All specifications include controls for invitations and invitations + comment form treatments, age and gender of respondent, mean adult education in the household, predicted household income, and dummies for type of household sampled.
* significant at 10%; ** significant at 5%; *** significant at 1%

## Table 7: Participation – First stage

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
|  | Attendance | Attendance of Non-Elite | Number Talk | Number Non-Elite who Talk |
| Invitations | 14.83*** | 13.47*** | 0.743*** | 0.286*** |
|  | (1.42) | (1.31) | (0.196) | (0.082) |
| Invitations + Comment | 11.48*** | 10.28*** | 0.498*** | 0.221*** |
|  | (1.41) | (1.33) | (0.174) | (0.072) |
| Meeting #2 | -5.32*** | -4.00*** | 0.163 | 0.024 |
|  | (1.16) | (1.11) | (0.162) | (0.088) |
| Meeting #3 | -4.29*** | -5.78*** | 0.431** | -0.158* |
|  | (1.26) | (1.19) | (0.180) | (0.093) |
| Stratum Fixed Effects | YES | YES | YES | YES |
| Observations | 1775 | 1775 | 1775 | 1775 |
| R-squared | 0.39 | 0.38 | 0.47 | 0.28 |
| Mean dep. variable | 47.99 | 24.15 | 8.02 | 0.94 |
| P-value Invitations = Invitations + Comment Forms | 0.04 | 0.04 | 0.23 | 0.45 |

Notes: Results from estimating equation (1), with the dependent variables the participation variables shown in the first row. Data is from the meeting survey. Each observation is a single village meeting. Includes stratum (subdistrict) fixed effects; since audit is constant within substrict, the audit variable is automatically captured by the stratum fixed effect. Robust standard errors in parentheses, adjusted for clustering at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

## Table 8: Participation – Impact on Meetings

|  | (1) | (2) | (3) |
|---|---|---|---|
|  | Number problems | Any corruption-related problem | Serious response taken |
| Invitations | 0.072 | 0.027** | -0.003 |
|  | (0.066) | (0.013) | (0.008) |
| Invitations + Comment | 0.104 | 0.026** | 0.015* |
|  | (0.067) | (0.013) | (0.008) |
| Meeting #2 | -0.187*** | 0.002 | -0.020** |
|  | (0.069) | (0.013) | (0.010) |
| Meeting #3 | -0.428*** | -0.036*** | -0.029*** |
|  | (0.077) | (0.012) | (0.010) |
| Stratum fixed effects | YES | YES | YES |
| Observations | 1783 | 1783 | 1783 |
| R-squared | 0.50 | 0.31 | 0.22 |
| Mean dep. variable | 1.18 | 0.07 | 0.03 |
| P-value Invitations = Invitations + Comment Forms | 0.62 | 0.96 | 0.02 |

Notes: Results from estimating equation (1), with the dependent variables the outcome of meetings shown in the first-row. Data is from meeting survey. 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 for an internal village investigation, asking for help from district project officials, or requesting an external audit. Estimation is by OLS. Robust standard errors in parentheses, adjusted for clustering by village.
* significant at 10%; ** significant at 5%; *** significant at 1%

## Table 9: Participation -- Main theft results

Panel A: Invitations

| | | | No Fixed Effects | | Engineer Fixed Effects | | Stratum Fixed Effects | | |
|---|---|---|---|---|---|---|---|---|---|
| *Percent missing:* *Log reported value –* *Log actual value* | Control Mean | Treatment Mean: Invites | Invite Effect | P-Value | Invite Effect | P-Value | Invite Effect | P-Value | Num Obs |
| Major items in roads | 0.252 (0.033) | 0.230 (0.033) | -0.021 (0.035) | 0.556 | -0.030 (0.035) | 0.391 | -0.026 (0.034) | 0.448 | 477 |
| Major items in roads and ancillary projects | 0.268 (0.031) | 0.236 (0.031) | -0.030 (0.032) | 0.360 | -0.032 (0.032) | 0.324 | -0.029 (0.032) | 0.356 | 538 |
| Breakdown of roads: | | | | | | | | | |
|   Materials | 0.209 (0.041) | 0.221 (0.041) | 0.014 (0.038) | 0.725 | 0.008 (0.038) | 0.841 | 0.005 (0.037) | 0.882 | 477 |
|   Unskilled labor | 0.369 (0.077) | 0.180 (0.077) | -0.187 (0.098) | 0.058 | -0.215 (0.095) | 0.026 | -0.143 (0.086) | 0.098 | 426 |

Panel B: Invitations + Comments

| | | | No Fixed Effects | | Engineer Fixed Effects | | Stratum Fixed Effects | | |
|---|---|---|---|---|---|---|---|---|---|
| *Percent missing:* *Log reported value –* *Log actual value* | Control Mean | Treatment Mean: Invites + Comment | Invite + Comment Effect | P-Value | Invite + Comment Effect | P-Value | Invite + Comment Effect | P-Value | Num Obs |
| Major items in roads | 0.252 (0.033) | 0.228 (0.026) | -0.022 (0.030) | 0.455 | -0.024 (0.029) | 0.416 | -0.015 (0.030) | 0.601 | 477 |
| Major items in roads and ancillary projects | 0.268 (0.031) | 0.238 (0.026) | -0.026 (0.032) | 0.409 | -0.025 (0.031) | 0.411 | -0.027 (0.031) | 0.385 | 538 |
| Breakdown of roads: | | | | | | | | | |
|   Materials | 0.209 (0.041) | 0.180 (0.032) | -0.028 (0.034) | 0.414 | -0.022 (0.033) | 0.501 | -0.010 (0.033) | 0.754 | 477 |
|   Unskilled labor | 0.369 (0.077) | 0.267 (0.073) | -0.099 (0.087) | 0.255 | -0.132 (0.088) | 0.136 | -0.090 (0.091) | 0.323 | 426 |

Notes: See Notes to Table 3. Results from estimating equation (1), a regression of the dependent variable on a dummy for audit treatment, invitations treatment and invitations + comment forms treatments. Each 'invitations effect' and 'invitations + comment effect' is from a separate regression, with the dependent variable listed in the row and the fixed effects specification listed in the column heading. Robust standard errors in parentheses. Regressions without stratum (i.e., subdistrict) fixed effects include a variable for audits, and allow for clustering of standard errors by subdistrict.

## Table 10: Net benefits calculation

| | Audits | | Invitations | |
|---|---|---|---|---|
| | Equal Weighted Net Benefits | Distribution Weighted Net Benefits | Equal Weighted Net Benefits | Distribution Weighted Net Benefits |
| Cost of treatment: | | | | |
| Monetary cost | -335 | -278 | -13 | -11 |
| Associated dead-weight loss | -134 | -111 | -5 | -4 |
| Time cost | -31 | -31 | -27 | -27 |
| *Subtotal* | -500 | -419 | -46 | -43 |
| | | | | |
| Change in rents received by corrupt officials: | | | | |
| From theft of materials | -367 | -224 | 62 | 38 |
| From theft of wages | -102 | -62 | -254 | -155 |
| Subtotal | -468 | -286 | -191 | -117 |
| | | | | |
| Change in benefits from project: | | | | |
| NPV value of road expenditures | 1165 | 1165 | 30 | 30 |
| Wages received by workers | 86 | 111 | 93 | 120 |
| Other expenditures | -37 | -37 | 73 | 73 |
| Subtotal | 1213 | 1238 | 196 | 223 |
| | | | | |
| **TOTAL NET BENEFITS** | **245** | **533** | **-41** | **64** |

Note: All figures in USD. Costs are listed as negative numbers. Distributional weights calculated using CRRA utility with coefficient of relative-risk aversion of 2. Derivation of change in NPV value of road and additional assumptions are discussed in the text.


## Table 11: Assumptions for loss ratios and worker capacity

| | Results from Calibration |
|---|---|
| *Loss Ratios* | |
| Sand | 1.00 |
| Rocks | 1.20 |
| Gravel | 1.75 |
| | |
| *Worker Capacity* | |
| Clearing brush and cleaning road surface (m2) | 20 |
| Spreading sand (m3) | 4.5 |
| Splitting rocks (m3) | 3.0 |
| Installing rock layer (m2) | 6.5 |
| Spreading gravel (m3) | 2.25 |
| Digging side channels and creating shoulder (m3) | 1.0 |
| Building retaining wall (m3) (unskilled labor portion) | 0.33* |
| Standard cut and fill(m3) | 2* |

Notes: All loss ratios are expressed as the ratio of the original amount of material to the amount measured, and are for measurements loose, not compacted. Worker capacity is the quantity of the given activity that can be done by one person per 6 hours of work.
* Original KDP assumption that was not able to be reconfirmed from calibration exercises.