USAID
FROM THE AMERICAN PEOPLE

ENDLINE IMPACT EVALUATION

GHANA STRENGTHENING ACCOUNTABILITY MECHANISMS (GSAM)

February 2018

This publication was prepared independently by Erik Wibbels (Duke University and Social Impact), Heather Huntington (University of Michigan and Cloudburst Group), Brett Gall (Duke University and Social Impact), and Elaine K. Denny (Duke University and UC Merced). It was produced at the request of the United States Agency for International Development as part of the Democracy, Human Rights and Governance – Learning, Evaluation and Research activity.

ENDLINE IMPACT EVALUATION

GHANA STRENGTHENING ACCOUNTABILITY MECHANISMS (GSAM)

February 2018

AID-OAA-M-13-00011

ABSTRACT

USAID/Ghana contracted Social Impact, Inc. to conduct an impact evaluation of USAID's Ghana Strengthening Accountability Mechanisms (GSAM) program, which aims to increase accountability of local District Assemblies in Ghana. This randomized-controlled trial, impact evaluation tests the effect of two distinct efforts to increase accountability and improve service delivery outcomes at the district level. One-hundred and fifty of Ghana's districts were matched and randomized into one of three groups: a top-down treatment group that received performance audits conducted by the central government Ghana Audit Service (GAS); a bottom-up treatment group that received civil-society organization (CSO) led scorecard campaigns; and a control group that did not receive either intervention. Through surveys with citizens, local administrators, and local politicians and through a review of administrative data, we find that both CSO and GAS programming generally reduce citizen satisfaction with projects and services, but this is largely driven by districts that receive negative audit reports. That citizens are correctly attributing bad audit performance to poor-performing DAs is encouraging from the point of view of accountability. This progress with citizens has not, however, translated into many substantial changes in how administrators or politicians manage projects or project budgets. Neither GAS nor CSO programming improve transparency or corruption. GAS programming does reduce the incidence of partisan manipulation of public resources by politicians, and it also increases the perception of partisan manipulation among administrators. This is consistent with GAS sensitizing administrators to partisan manipulation and reducing its actual incidence among DA politicians. CSO programming increases citizen-reported consultation on recent development, and administrators in CSO districts spend, on average, three hours more responding to constituents. Reasons that the intervention did not have a stronger impact on district officials includes (1) natural limits to the number of citizens reached by the intervention, (2) limited district government capacity and frequent turnover, and (3) local government dependence on federal budget transfers.

ACKNOWLEDGEMENTS

The authors would like to thank the following people for their generous help: Emmanuel Mensah-Ackman and Steve Edminster of USAID/Ghana; Peter Quartey, Joseph Teye, George Owusu, and all of the great field staff at ISSER; Jacob Essilfie and the rest of the team at the Ghana Audit Service; Clement Tandoh, and the entire team at CARE International/Ghana and partnering organizations; Adjei Roberson of the Ghana Statistical Service; Clement Tandoh, and the entire team at CARE International/Ghana and partnering organizations; Morgan Holmes of USAID's Center of Excellence for Democracy, Human Rights and Governance; Alison Miranda and Daniel Sabet of Social Impact; and last but certainly not least, Benjamin Kauffeld, the former DG officer at USAID/Ghana. This evaluation would not have been possible without this great group of people.

Wibbels would like to dedicate this report in fond memory of Sarkodie Gyekye, the head of evaluation for CARE International's Ghana office, who was so generous with his time and friendship.

CONTENTS

TABLES

FIGURES

ACRONYMS

APR	Annual Progress Report
ARIC	Audit Report Implementation Committee
ATE	Average Treatment Effect
CATE	Conditional Average Treatment Effect
CBO	Community Based Organizations
CO	Contracts Officer
COP	Chief of Party
COR	Contracts Officer's Representative
CSO	Civil Society Organization
DA	District Assembly
DACF	District Assemblies' Common Fund
DCE	District Chief Executive
DDF	District Development Fund
EA	Enumeration Area
FOAT	Functional and Organizational Assessment Tool
GAS	Ghana Audit Service
GSAM	Ghana Strengthening Accountability Mechanisms
GSS	Ghana Statistical Service
IE	Impact Evaluation
IGF	Internally Generated Funds
IRB	Institutional Review Board
ITT	Intent to Treat
M&E	Monitoring and Evaluation
MDES	Minimum Detectable Effect Size
MMDA	Metropolitan, Municipal, or District Assembly
OLS	Ordinary Least Squares
PAC	Public Accounts Committee
PFM	Public Financial Management
PM	Presiding Member

RCT	Randomized Control Trial
SSA	Sub-Saharan Africa
TOT	Treatment on the Treated
USAID	United States Agency for International Development
USG	United States Government
VSLA	Village Saving and Loan Associations

EXECUTIVE SUMMARY

INTRODUCTION

In 1988, Ghana launched a comprehensive local government and public administration reform process based on six pillars of decentralization: political, administrative, fiscal, planning, market, and spatial. The reform sought to transfer power and competence to a local government of Metropolitan, Municipal or District Assemblies (MMDAs), which are similar in scale to U.S. counties. Advocates of decentralization in Ghana have argued that it can promote participatory governance, improve service delivery, and facilitate socio-economic development (Ohene-Konadu 2001). Nonetheless, the MMDA system is failing to achieve its essential mandate of efficient, effective, and equitable local service delivery (Ayee and Dickovick 2010; CDD-Ghana 2008; Fox et al. 2010). Decentralization reforms have failed to address many of the basic problems associated with poor governance and service provision at the local level, and local officials appear insufficiently accountable to either the central government or to the citizens they represent.

This report presents the results of an impact evaluation (IE) of USAID's Ghana Strengthening Accountability Mechanisms (GSAM) program, which aims to strengthen unaccountable or weakly accountable local District Assemblies (DAs). The evaluation explores the impact of two approaches to increasing accountability: (1) a series of top-down performance audits by the central government's Ghana Audit Service (GAS) of District Assembly (DA)-contracted and constructed development projects and (2) a bottom-up civil society organization (CSO)-led campaign also focused on the quality of district development projects. Citizen outreach to disseminate findings pertaining to each respective district's programming has occurred in both the top-down and bottom-up audits districts. Both the interventions focus on auditing districts' capital development projects, including schools, public toilets, clinics, feeder roads, irrigation projects, and other infrastructure and centers of public service provision. These projects represent the majority of district government spending.

EVALUATION DESIGN

The evaluation rigorously assesses the impact of performance audits and CSO-led social audits on improving accountability and service delivery outcomes through a randomized control trial. One hundred fifty districts were randomly assigned to three treatment conditions. Fifty districts received performance audits from the GAS and subsequent dissemination of the performance audit results through meetings with district officials and a public scorecard campaign. Fifty additional districts received social audits from civil society organizations and subsequent dissemination of the social audit results through a similar scorecard-based information dissemination campaign. Another 50 districts serve as a pure control group and did not receive audits or any information dissemination campaign. As designed, this is one of the largest IEs ever conducted on a USAID governance project at the district level and one of the first to explicitly compare top-down and bottom-up interventions for promoting accountability in district governance.

The primary outcomes of interest include (1) the quality of district development projects and corresponding service access, (2) the quality of district governance, with special reference to district capital projects, and (3) accountability of district officials and citizen engagement with district-level democratic procedures. These outcomes are measured across through three surveys and an analysis of administrative data. The surveys target three key populations: citizens, who GSAM aims to empower to be better-informed participants in district governance; district politicians, who make decisions on which development projects to build in which

communities; and district administrators, who are responsible for overseeing the contracting and construction of development projects. In addition, the evaluation team digitized project completion rates from the 2016 district *Annual Progress Reports* and the results of the Ghana Audit Service's annual report on district management of public funds in 2016 (GAS 2017). Baseline and endline data was collected across these data sources; however, given the targeted nature of capital development projects additional citizen surveys were conducted in high-intensity areas, i.e. communities where GSAM was most active. Because of good balance in baseline data and some important differences between baseline and endline data collection, the primary analysis relies exclusively on endline data. The research plan was preregistered with EGAP (Huntington, Wibbels, and Gall 2017).

RESULTS

We hypothesized that GSAM programming (both CSO and GAS interventions) would have the following effects:

1. Improvement in the quality of capital projects and service delivery provided at the district level
2. Improvement in the quality of district governance
3. Increased government accountability and citizen engagement

We test each of these three main hypotheses in multiple ways. The text below provides a summary of the conclusions across the study's three main hypotheses. Conclusions related to the citizen survey are drawn from results in high-intensity GSAM communities. As noted in the main report, results are considerably more modest outside these high-intensity GSAM communities.

In summary, while we find null effects on many outcome variables, there are also some important positive effects. GSAM's CSO programming has generally had more effects on citizen engagement than GAS programming, while GAS programming has had a bigger effect on governance quality. Across treatment arms, program effects have generally been stronger on citizen perceptions of district governments and their stated willingness to hold DA officials accountable, even if that has not been accompanied with more activism or participation. GSAM has limited effects specifically on the governance of development projects, but GAS programming has encouraging effects on overall governance quality, particularly the incidence of partisan manipulation of public resources. The highlights bearing on each hypothesis are:

- **Quality of projects and services:** Both CSO and GAS programming generally reduce citizen satisfaction with projects and services, but this is largely driven by districts that receive negative audit reports. That citizens are correctly attributing bad audit performance to poor-performing DAs is encouraging from the point of view of accountability. This progress with citizens has not, however, translated into many substantial changes in how administrators or politicians manage projects or project budgets.
- **Quality of governance:** Neither GAS nor CSO programming improve transparency or corruption. GAS programming does reduce the incidence of partisan manipulation of public resources by politicians, and it also increases the perception of partisan manipulation among administrators. This is consistent with GAS sensitizing administrators to partisan manipulation and reducing its actual incidence among DA politicians.
- **Accountability:** GSAM has no effect on citizen participation in district governance, but both GAS and CSO programming increase the likelihood that citizens will hold DA officials accountable for project quality and are more likely to turn to a DA official for help with projects. CSO programming increases citizen-reported consultation on recent development, and administrators in CSO districts spend, on average, three hours more responding to constituents.

As discussed below, changing deep institutional practices bearing on project completion rates and financial irregularities will likely require more sustained institutional and fiscal change than is possible with a program like

GSAM. Nevertheless, the overall results are consistent with some important improvements in accountability and project outcomes—citizens are sensitive to negative project outcomes and report increased willingness to hold DAs accountable. At the same time, DAs have somewhat reduced partisan manipulation of budgets and increased responsiveness to citizens. In short, GSAM has had some positive effects on the citizen-led, demand side of accountability, and some on the supply side of governance. The findings beg two follow-up questions:

Given that many impact evaluations have found bottom-up accountability efforts to be mixed, why do we see stronger effects on citizens than some other, related impact evaluations? While other accountability-oriented governance projects focus on features of politics that are somewhat remote from citizens' day-to-day lives (such as, for instance, how elected officials spend their time), GSAM focuses on development projects that have very important, immediate implications for citizen wellbeing. Rural citizens care very much if they have a local school, a public toilet or a functional market.

Why are the effects on administrators and political officials more modest? The weakness of the findings among DA officials likely reflects several factors. First, GSAM has still reached a relatively small share of the broader DA citizenry. Thus, any pressure coming from the ground up is likely to be modest. Second, the underlying institutional capacity of DAs is limited by frequent administrator rotation, low reelection rates among political officials and scarce resources and skills for administering large development projects. These limitations are both political and administrative. Perhaps most importantly, they limit the incentives and capacity of DA officials to learn from, and respond to, citizens newly empowered by GSAM. Third, the system of fiscal transfers in Ghana makes DA budgeting and contracting difficult. Since central transfers are often delayed, contractors sometimes go unpaid, and projects are left incomplete. To the extent DA officials cannot impact this binding fiscal constraint, this also limits their incentives to invest in being more responsive to citizens or institutional innovations that might improve DA institutional planning and oversight of projects.

POLICY RECOMMENDATIONS

This overarching interpretation of GSAM's effects suggest a number of policy recommendations.

1. GSAM should further expand the extent of citizen outreach in CSO districts.
2. USAID and GSAM should develop and implement programming to strengthen the internal institutional capability of DAs. This should include development of an internal monitoring and evaluation system for development projects, training for key DA members and administrators involved in project planning and oversight, and efforts to professionalize the Audit Report Implementation Committees (ARICs).
3. USAID, the GAS, and the government of Ghana (GoG) should explore and experiment with additional methods to increase performance audit impact based on the experiences of other countries.
4. The GoG should improve the regularity of central transfers from Accra to the districts.
5. The GoG should reduce the rotation of district administrators.
6. The GoG should consider efforts to professionalize DAs and make careers as DA politicians more attractive.

FUTURE POLICY-RELEVANT RESEARCH RECOMMENDATIONS

GSAM has pointed to a number of opportunities for future research on district governance. We briefly emphasize four of them here.

1. USAID and GSAM stakeholders should continue to assess the medium- to long-term impact of GSAM with an additional round of data collection.

2. USAID and GSAM stakeholders should tests how best to leverage performance audit capacity, exploring variation in the frequency and nature of performance audits, complementary programming (e.g., training, follow-up policies, sanctioning/promotion), and dissemination.
3. Given the dependence of reform on local officials and the current high turnover rates, future research should seek to gain a better understanding of district-level political and administrative careers.
4. Given profound inequality in access to development projects communities, future research should also assess why district governments allocate development projects to some communities but not others.

INTRODUCTION

The USAID Ghana Strengthening Accountability Mechanisms (GSAM) program is a five -year intervention designed to strengthen unaccountable or weakly accountable District Assemblies (DAs). The GSAM program represents a ten million-dollar project that is supported by USAID/Ghana's Democracy and Governance Office. It is being implemented across 100 districts in Ghana, from November 2014 through the end of 2019. The objective of the program is to improve local governance in Ghana through strengthening both top-down and bottom-up accountability. The work addresses USAID/Ghana's objectives of improving governance, improving the quality of district-built capital projects and related service delivery, building the capacity of Ghanaian institutions, and strengthening local democratic procedures and citizen engagement.

This report presents the results of an impact evaluation (IE) of the GSAM program. The objective of the evaluation is to determine the impact of two GSAM interventions: (1) a series of top-down performance audits by the central government's Ghana Audit Service (GAS) of District Assembly (DA)-contracted and constructed development projects and (2) a bottom-up civil society organization (CSO)-led campaign also focused on the quality of district development projects. Citizen outreach to disseminate findings pertaining to each respective district's programming has occurred in both the top-down and bottom-up audits districts.

Both the top-down and the bottom-up interventions focus on auditing districts' capital development projects. These projects include schools, public toilets, clinics, feeder roads, irrigation projects, and other infrastructure and centers of public service provision, and they represent the majority of district government spending. Throughout this report we refer to these district-built capital projects as "development projects."

When these district development projects are planned and executed with citizen needs in mind, they can impart substantial developmental benefits for citizens. The government of Ghana, USAID, and citizens all expressed concern that many district governments were building projects inefficiently, ineffectively, and without citizen input. For example, there is evidence that more than 40 percent of projects are never completed (Williams 2016). In light of these concerns, the GSAM program was designed to increase accountability for the planning and execution of these crucial district government projects.

The evaluation rigorously assesses the impact of performance audits and CSO-led social audits on improving accountability and service delivery outcomes through a randomized control trial. One hundred fifty districts were randomly assigned to three treatment conditions: GAS audits, CSO audits, a control group. Specifically, 50 districts received performance audits from the GAS and subsequent dissemination of the performance audit results through meetings with district officials and a public scorecard campaign. Fifty additional districts received social audits from civil society organizations and subsequent dissemination of the social audit results through a similar scorecard-based information dissemination campaign. Another 50 districts serve as a pure control group and did not receive audits or any information dissemination campaign. As designed, this is one of the largest IEs ever conducted on a USAID governance project at the district level and one of the first to explicitly compare top-down and bottom-up interventions for promoting accountability in district governance.

Prior to the rollout of project activities, baseline data collection was carried out in each of the 150 districts. This data collection effort sought to gather information on baseline conditions on the key outcomes of interest. The endline data collection, the primary data source for this endline evaluation, is also focused on similar outcomes. These primary outcomes of interest in the endline evaluation include:

1. The quality of district development projects and corresponding service access.
2. The quality of district governance, with special reference to district capital projects
3. Accountability of district officials and citizen engagement with district-level democratic procedures
4. The objective performance of districts in completing development projects and managing budgets
5. Awareness of GSAM and its impact on citizen information and engagement with district government

These outcomes are measured across three surveys and an analysis of administrative data. The surveys target three key populations: citizens, who GSAM aims to empower to be better-informed participants in district governance; district politicians, who make decisions on which development projects to build in which communities; and district administrators, who are responsible for overseeing the contracting and construction of development projects. In addition, the evaluation team digitized project completion rates from the 2016 district *Annual Progress Reports* and the results of the Ghana Audit Service's annual report on district management of public funds in 2016 (GAS 2017). Baseline and endline data was collected across these data sources; however, given the targeted nature of capital development projects additional citizen surveys were conducted in high-intensity areas, i.e. communities where GSAM was most active. Because of good balance in baseline data and some important differences between baseline and endline data collection, the primary analysis relies exclusively on endline data. The research plan was preregistered with EGAP (Huntington, Wibbels, and Gall 2017).

1.0 BACKGROUND

In 1988, Ghana launched a comprehensive local government and public administration reform process based on six pillars of decentralization: political, administrative, fiscal, planning, market, and spatial. The reform sought to transfer power and competence to a local government of Metropolitan, Municipal or District Assemblies (MMDAs), which are similar in scale to U.S. counties. Within the MMDAs, elected District Assembly (DA) members are elected to represent local citizens, and a District Chief Executive (DCE) is appointed by the President of the Republic to represent central authorities. Although the MMDAs are nominally non-partisan, with candidates not promoted by political parties, presidents have historically appointed members from their own parties (Ayee and Dickovick 2010). As of June 2012, there were 216 MMDAs in Ghana.

In Ghana, MMDAs are the only substantively meaningful level of government below the national administration. The MMDAs are primarily responsible for improving service delivery outcomes through "local-level policy and planning, [which is] based on, and informed by, national policies and programs." (Ayee and Dickovick 2010, pg. 135) MMDAs rely on several sources of revenue to fund local government policies and activities. The District Assemblies' Common Fund (DACF) provides inter-governmental transfers to MMDAs from the central government. DACF serves as the MMDA's most important source of revenue, and it is distributed according to a revenue sharing formula. Additionally, Internally Generated Funds (IGFs) are derived from taxes and levies imposed by MMDAs. Further sources of revenue include the District Development Fund (DDF) and rents from local extractive industries (Ayee and Dickovick 2010).

Advocates of decentralization in Ghana have argued that it can promote participatory governance, improve service delivery, and facilitate socio-economic development (Ohene-Konadu 2001). Consistent with decentralization initiatives in other countries, reforms have been guided by the belief that local governments are better able to respond to citizen needs and are more accountable to citizens due to their proximity to the citizenry. The expectation has been that well-managed local governments will be an important means to achieve democracy, "good governance," and the efficient mobilization and allocation of scarce resources to meet citizens' needs.

Ghana's ambitious decentralization reform process and various donor initiatives notwithstanding, the MMDA system is failing to achieve its essential mandate of efficient, effective, and equitable local service delivery (Ayee and Dickovick 2010; CDD-Ghana 2008; Fox et al. 2010). Decentralization reforms have failed to address many of the basic problems associated with poor governance and service provision at the local level. Corruption, rent-seeking, and low-effort service provision continue to characterize some local governments. These characteristics combine to facilitate local elite and interest group capture of decentralized politics, distort democratic processes, and obfuscate lines of accountability. In the absence of accountability, local elected officials and technocrats face weak incentives to address the needs of broad groups of local citizens. This has significant negative implications for the quality of local public services.

Some of the underlying challenges to accountability result from structural factors that are all but impossible to address with USAID programming. For instance, the fact that DCEs are very powerful and appointed by the president profoundly limits their incentives to respond to district citizen concerns and receive input from citizens. Since DCE promotion involves returning to the national government, they are incentivized to respond to the national party leadership over local needs. Due to these institutional incentives, any USAID or donor funded projects targeted at this group is unlikely to produce substantial improvements in district governance, accountability, and development projects.

Other gaps in local accountability, however, are amenable to improvement through USAID or other donor programming. GSAM focuses its efforts on two areas with considerable promise for improvement: (1) limited central government oversight of the MMDAs and (2) limited availability of information to citizens on MMDA budgets, development projects and service delivery.

Looking from the top-down, there are clear shortcomings with the existing means of central government monitoring of district governments' performance. Although the law requires that GAS conduct a financial audit of all MMDA financial accounts for each fiscal year, the system of financial auditing is in practice a limited tool for promoting accountability. First, the financial audits are only intended to identify accounting irregularities, and they do not provide information on the quality of roads, irrigation projects, or any other public expenditure. They also do not provide information on the efficiency or effectiveness of MMDA expenditures. Second, the sanctioning regime for poor audit performance is very weak. Audit results first go to the DCE and the District Assembly's Audit Report Implementation Committee (ARIC), which are free to ignore any violations. While all MMDA audits are later aggregated in a report that is delivered to the Auditor General and from there to the Parliament's Public Accounts Committee (PAC), the PAC does not have the capacity to formally sanction MMDAs despite evidence of corruption in public financial management (PFM). The PAC can rely on public shaming and the press has taken up the cause in a few cases, but only a small number of MMDA budget violations are ultimately sanctioned or face public scrutiny. A recent Supreme Court ruling, however, provides a glimmer of hope for success in the fight against corruption and impunity in District Assembly financial administration. In June 2017 the Supreme Court ordered the Auditor General to surcharge any official found to have misappropriated public funds.

Looking from the bottom up, local citizens also have limited means to hold local government officials accountable. Citizens possess limited information on MMDA budgets, execution of that budget (e.g., on development projects) or the relative quality of local public services that results. In many districts, the weakness of Assembly-level oversight of the budget further exacerbates this problem. Internal auditing procedures are compromised by weak capacity and poor system design. The Assemblies have ARICs, but heavy representation of DCE interests (including the participation of the DCE him/herself) and limited means of public outreach compromise the work of the ARIC. As such, citizens have few means to directly -- or indirectly through their elected officials -- assess budget expenditure or quality of district development projects. Moreover, since citizens typically do not observe the quality and efficiency of schools, health clinics or irrigation projects in other localities or districts, it is difficult for them to compare the performance of their local officials against the performance of other jurisdictions. All of these challenges were reflected in baseline household surveys, which showed limited engagement with, and considerable dissatisfaction with district governance.

Several development failures emerge from this lack of information and lack of accountability, including weak PFM, poor quality development projects, room for corruption, and low levels of citizen engagement. Knowledge and participation are crucial mechanisms for generating good governance and low levels of either threatens accountability of district governments.

Performance audits and social audits have the potential to fill this information gap and increase accountability. These interventions can provide information to citizens on the performance of their local governments, specifically with regard to the management of district development projects. These audits provide information on: (1) whether projects are planned in a participatory manner, (2) the degree of competitiveness and reliance on merit in project procurement, (3) whether the projects are completed, (4) whether project construction is well-executed to produce high-quality infrastructure, and (5) the extent to which the projects benefited local citizens.

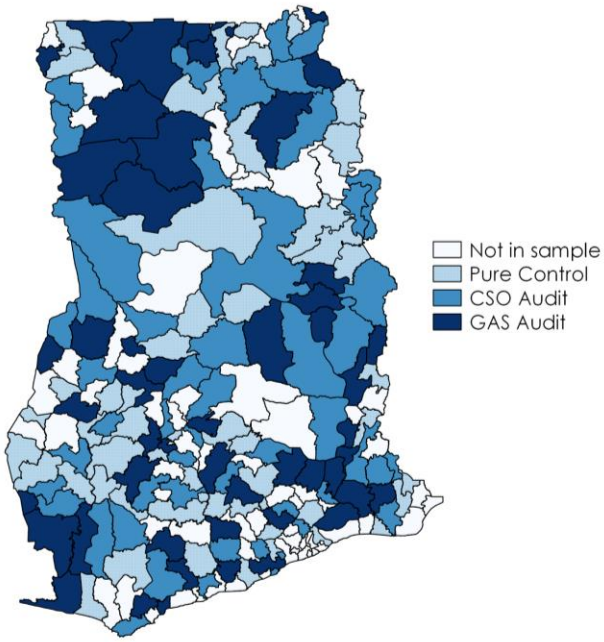
Research has shown powerful positive effects of performance audits on reducing local corruption (Olken 2007) and increasing local electoral accountability (Ferraz and Finan 2008). In addition, prominent recent randomized trials have also examined the effect of bottom-up social accountability mechanisms for the quality of local public services. Andrabi, Das, and Khwaja (2009), for instance, find that providing details on school performance has measurable effects on both learning and school fees. Duflo, Hanna, and Ryan (2012) and Muralidharan and Sundararaman (2011) show that even moderate improvements in teacher incentives can increase teacher attendance. Björkman and Svensson (2009) examine the effects of community-based monitoring on local health clinics in Uganda and report impressive improvements in service provision, utilization, and health outcomes. All of these interventions are aimed at improving social accountability by providing citizens with the information necessary to make better-informed demands on the government.

In summary, the overarching development problem that this program aims to address is that MMDA officials are unaccountable. They are insufficiently accountable to the central government, which funds the largest share of MMDA budgets. They are also unaccountable to local citizens, who are the most important beneficiaries of district-built development projects. Building on cutting edge development projects elsewhere, the impact evaluation of this program involves the randomized assignment of two main activities – central government performance audits and a citizen scorecard campaign – along with the randomization of control group status to a third group of districts. Given this design, the impact evaluation allows the IE team to assess the effectiveness of each activity.

2.0 INTERVENTIONS AND HYPOTHESES

The GSAM implementation began in 2014 and is expected to take place over a 5-year period, across a total of 150 districts in all ten regions of Ghana. The IE team developed the endline report prior to the end of the program in order to accommodate specific features of program timing. The two activities under evaluation include: (1) GAS Performance Audits and corresponding citizen information campaigns; and (2) a CSO-led social audit and information campaign. The impact of each of these activities are identified by a randomized control trial. Fifty districts received GAS performance audits and scorecard campaign, and 50 districts received the CSO-led social audits and scorecard campaign. An additional 50 districts serve as a control group during the period of the impact evaluation.

Figure 1: GSAM Experimental Conditions Assignment by District



2.1 INTERVENTION 1: GAS PERFORMANCE AUDIT & INFORMATION CAMPAIGN

As part of GSAM, the GAS implemented performance audits of the DAs to strengthen hierarchical accountability across levels of government. As noted above, although the GAS is legally mandated to conduct financial audits of all MMDAs, financial audits provide limited information on the quality of local governance because they do not address the quality of local public services or capital spending. Performance audits, on the other hand, are explicitly aimed at evaluating the value-for-money in public budgeting. They go well beyond accounting to cover each step in the process of planning and building development projects, including: (1) the identification and budgeting for projects in district budgets, (2) the tendering and award of contracts, (3) the monitoring,

evaluation, and payment of contractors as projects are being built, and (4) post-completion assessments of project quality and the impacts of projects on citizen welfare. Prior to 2014, the GAS had conducted performance audits of central ministries and some infrastructural projects; however, it had neither staff nor capacity to do so at the district level on a large-scale basis. With GSAM support, the GAS could develop this capacity and implement performance audits at the district level for fifty district assemblies.

GAS selected for auditing the top two most expensive school-building projects reported in Annual Progress Reports of each of the 50 districts assigned to the GAS audit experimental condition.[1] The choice was made to focus on schools as they are easily the most common district-built project, and focusing on schools allowed for a simplification of the training of auditors.[2] These projects include classroom block construction and other infrastructural projects.

The GAS submitted notifications to 50 DAs randomly selected by the evaluation team, including the DCE, members of the Assembly, the budget and development committees, and the District Coordinating Director — informing them that they would each receive a performance audit. The letter explained that performance audits are much more involved than standard GAS financial audits. A typical audit took two weeks and involved a combination of interviews, document review, and field work. The performance period reviewed covered a three-year period, 2012-2014. Following completion of the performance audits, GAS conducted exit meetings with district officials to provide preliminary feedback and findings. Thereafter, the GAS produced audit reports. All of the performance audits were completed by late spring 2015.

Thereafter, GAS and implementing partner CARE International disseminated results to DA officials and citizens. This began with the construction of district scorecards on the basis of the audit results (see Appendix 1). Through December 2015 and January 2016, the evaluation team worked closely with GAS and CARE International (the contractor for the dissemination of the information to district officials and citizens) on how to score the audit reports with an eye to the development of clear and accessible scorecards. This resulted in a sensible and simple way of scoring the reports. CARE and GAS scored the reports separately and then compared notes in order to ensure that the scores reflected GAS' intent. The resulting scorecards are quite straightforward and include both objective scores and information on how the district did compared to others. The scorecards include simple graphics for those citizens that are illiterate and/or innumerate.[3]

To present the findings and scorecards, GAS and CARE conducted widely publicized, public information-sharing sessions with DA officials and invited citizens and CSOs in all 50 district capitals during April and May of 2016. This timing proved a serious difficulty for the midline survey team (see below), but the GAS/CARE scorecard events appear to have gone well. It seems DA officials were appropriately concerned about the findings, and some citizen groups reported that this was the first time they had heard that they could complain about incomplete and poorly constructed projects.

Lastly, a large-scale public information and mobilization campaign followed the series of DA meetings in the GAS treated districts, which included: a) scorecard dissemination events in 25 communities per GAS district; b) dissemination events in "alternative spaces" to reach individuals who are often excluded from formal events, including youth, women and those with disabilities; and c) radio jingles to distribute audit results and encourage

[1] These reports are required and their filing is a precondition for receiving the following year's fiscal transfers from the central government. They include information on each project, including project type, location, funding source and level of completion.
[2] The inclusion of water or electrification projects, for instance, have quite different engineering challenges; thus, assessments of project quality are quite different than for buildings like schools.
[3] See Annex 2 for an example of the scorecards.

project monitoring. Scorecards were distributed by CARE International and their partners to DA citizens at public events in each of the 50 GAS audit districts, resulting in 1250 meetings.[4] In addition to the distribution of scorecards, the meetings also involved the provision of information regarding the DA, including the names of DA officials responsible for project quality, members of the relevant DA committees, and the district administrators. CSOs also engaged with DA citizens to demand accountability for the results. These 1,250 local meetings were facilitated by the coordinating CSOs and district-level community-based organizations (CBOs), and they provided a forum for discussing the scorecards and mobilizing citizens around the results.

The dissemination events that occurred in "alternative spaces" were aimed at typically marginalized populations and took place in schools, mosques, churches, markets, and taxi stands. These events covered similar material to that discussed above; there were 123 such events reaching 12,099 citizens across the 50 GAS districts. Finally, the radio jingles aired GAS audit results, encouraged citizen participation in oversight of development projects, and provided phone-in segments for public discussion.[5]

The GAS program activities could have influenced DA planning, budgeting, and project oversight through two mechanisms. First, the performance audits themselves could have promoted changes in district practices. Indeed, GAS, CARE, and our own team noted that DA officials were nervous about the arrival of the GAS in their districts and about the content of the initial reports, as conveyed in the exit interviews. Second, it could be that citizen attention to audit results is a prerequisite for behavioral change in the DA. In those DAs that receive positive reports, it is expected that the campaign will improve citizen attitudes about the quality of local governance. In those DAs that receive negative reports, it is expected that citizens will demand improved projects and services and demand accountability.[6] Regardless of the outcome of the GAS audit, with improved information, both central government officials and citizens should be empowered to hold district governments more accountable for their management of the public purse.

2.2 INTERVENTION 2: CSO-LED SOCIAL AUDIT & INFORMATION CAMPAIGN

The CSO-led social audit (SA) campaign was implemented in an additional 50 districts randomly selected by the evaluation team. The CSO social audit was designed to strengthen social accountability by providing citizens with detailed information - collected via the social audits - on the quality of district capital projects. This portion of the programming was contracted to CARE International. According to project reporting, CARE led a consortium of implementing partners built on its preexisting network of 245 CSOs, over 10,388 farmers' groups, youth groups, and 2,400 village saving and loan associations (VSLAs). These networks and relationships provided the on-the-ground infrastructure for sustained social audits involving community-gathered information on the planning, passage, implementation, and use of district development budgets.

The social audit programming has four key components: First, CARE and its partners formed District Steering Committees, typically with seven members, who are responsible for overseeing GSAM activities in the district, promoting citizen participation in project budgeting and monitoring, and engaging with DA officials on capital budgeting and project oversight. Second, as in the GAS districts, two capital projects per district were selected from the 2015/16 Annual Development Plans for extensive auditing. Third, CARE International and CSO partners

[4] 1,185 of these meetings had occurred by the end of the third quarter of 2016; 53 additional meetings occurred in the fourth quarter of 2016; the final 12 meetings occurred in the first quarter of 2017.
[5] As of March 2017, this radio-based dissemination had occurred in 45 of 50 GAS districts. The five where they had not occurred are Biakoye, South Dayi, Central Tongu, Asuogyaman, and North Tongu.
[6] Given the relatively small number of districts involved in this arm of programming, our capacity to detect these heterogeneous effects is limited; we discuss this challenge in greater detail in the methodology section.

trained community members as Community Development Monitors (CDMs) to monitor work quality, absenteeism by contract, the pace of work, and the quality of work materials. They were also provided names and phone numbers and instructed on procedures for registering complaints with DA officials about projects. These CDMs then worked with CSOs in auditing projects on a biweekly basis.[7] Fourth, the results of these social audits were aggregated into scorecards modeled on those developed for the GAS audits. The construction of the scorecards resulted from a combination of community meetings (which produced an initial "community scorecard") and CSO monitoring data. Fifth, the resulting citizen scorecard was disseminated across the district in similar ways to those described above in the GAS districts, albeit less extensively. As in the GAS districts, dissemination has occurred via 50 district-level town hall meetings in all CSO districts), 676 community meetings across the CSO districts (as compared to 1250 in GAS districts), and community radio/jingles.[8] Unlike in the GAS districts, social audit results were also disseminated via billboards in 47 of the 50 CSO districts. Dissemination events did not take place in "alternative spaces" in the CSO districts.

The local training for the social audits was scheduled to begin in the second quarter of 2015. Due to the difficulties in building and coordinating CSO capacity in so many districts, the programming was delayed until the last quarter of 2015. This delay had two significant implications for the impact evaluation. First, it made a midline evaluation of the CSO-led social audit campaign moot since too few activities had taken place. Second, the delay placed considerable time pressure on the CSOs to complete their social audit activities so that the requisite information would be in the hands of citizens prior to the annual DA budgeting processes.

Despite these difficulties, brief fieldwork by the evaluation team in three districts in March 2016 and subsequent quarterly reports by CARE International indicate that the CSO programming met most of its targets before the endline data collection went into the field. Social audits took place of all 100 projects in the 50 CSO districts, and CARE and partners generated scorecards for all 50 districts. District-level town hall meetings have taken place in all 50 CSO districts, 46 of which occurred before endline data collection began. Community-level dissemination was less widespread than in GAS districts, but reached 27,278 citizens in 613 communities across the 50 CSO districts; most of these community meetings occurred in the same quarter as endline data collection, so it is unclear if the evaluation team captured their effect. Likewise, the construction of billboards occurred during the quarter of data collection.

2.3 THEORY OF CHANGE AND ASSUMPTIONS

The GSAM Theory of Change (ToC) is presented below and links the program activities to the expected outcomes. The key risks and assumptions are included along the results chain in blue circles. These risks and the way that they potentially complicate the ToC are discussed in-depth in Section 3.4 below as part of the endline challenges and cautions. For both the CSO and GAS interventions, audits (social or performance) provide improved information to citizens regarding the quality of district capital projects and service delivery. The dissemination of that information is expected to increase citizen empowerment and citizen engagement with district-level politics and democratic procedures. This in-turn promotes greater DA accountability, which should simultaneously feed into improved district governance and improved DA budget and development project performance. Independent from the information campaign, performance audits by GAS are anticipated to promote hierarchical accountability and improvements on DA planning, budgeting and project oversight. The first stage outcomes of interest are increased quality of district development projects and improved service delivery. The ultimate long-term outcome of interest is improved livelihoods and welfare for citizens.

[7] The project completion rates in these CSO communities was 56%, which compares favorably to the rate of 45% in 2015 for all districts.
[8] Nine CSO districts had not received this radio programming by March 2017.

Figure 2: CSO Causal theory, with Assumptions Numbered in the Blue Circles

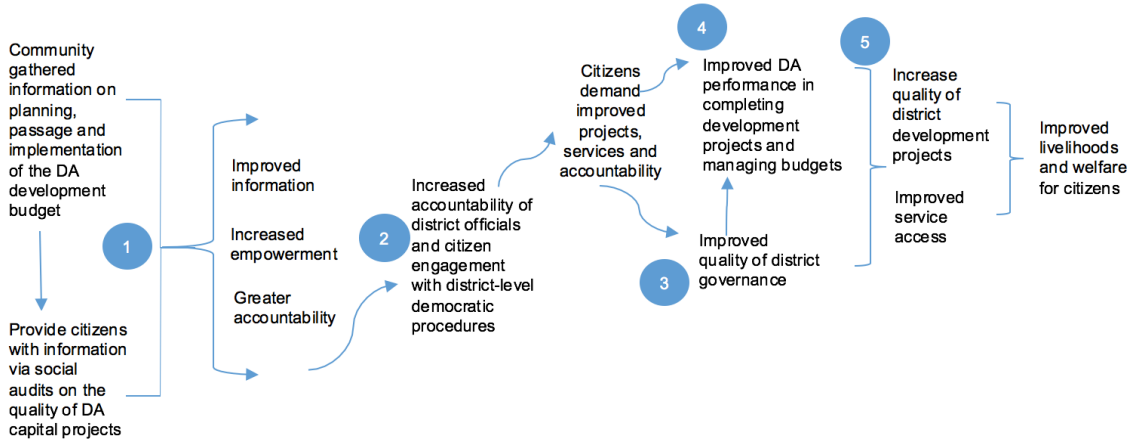
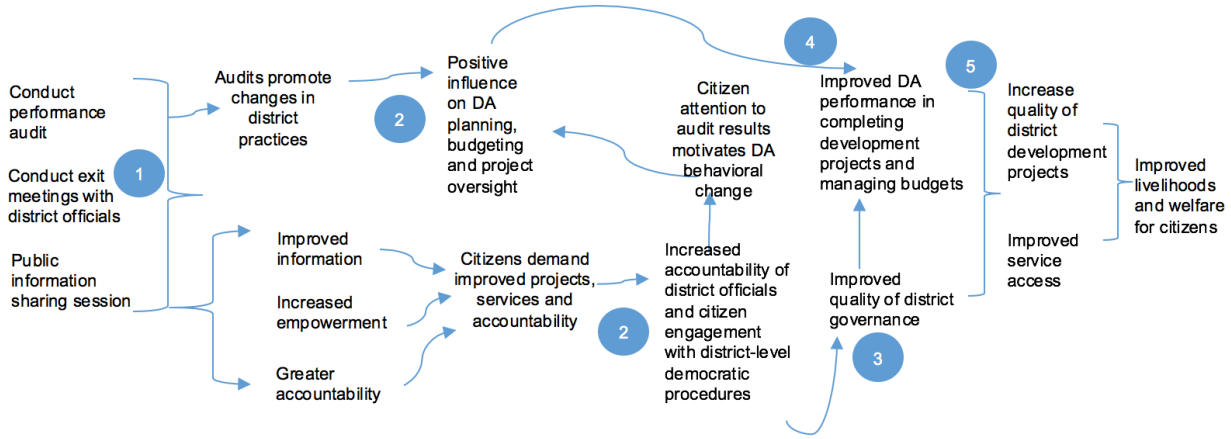


Figure 3: GAS Causal Theory, with Assumptions Numbered in the Blue Circles



GSAM Key Risks and Assumptions

1. Activities are implemented effectively according to the intervention plan. GAS/CARE reduce the information gap through a comprehensive public dissemination effort; a majority of citizens are reached in programmed districts.

2. Information needs to be in citizen's hands prior to DA budgeting processes.

3. The timing of district and national elections needs to be taken into consideration, as these have an impact on the composition of the DA government.

4. Delays in the passage of the DA budgets complicates the ability for citizen pressure to be brought to bear on the development of the DA Annual Action Plan.

5. Sufficient time is needed to achieve higher level outcomes.

2.4 HYPOTHESES

This evaluation tests a series of hypotheses, which we organize disaggregate by measurable outcomes and target population.

Hypothesis 1: GSAM improved capital projects and service delivery

For citizens, the interventions will:

- Increase access to development projects and the corresponding services,
- Reduce obstacles to completing projects, and
- Increase citizen satisfaction with services.

Within ***district governments*** (i.e. among administrators and political officials), the interventions will:

- Improve district planning
- Improve district contracting, and
- Improve oversight of development projects.

Hypothesis 2: GSAM improved governance quality

The interventions will:

- Improve perceptions of the relative performance of district governments
- Reduce the incidence of corruption, and
- Increase government transparency.

Hypothesis 3: GSAM improved district government accountability and citizen engagement

The interventions will:

- Improve the responsiveness of DAs to community needs when it comes to both budgeting and locating development projects.
- Increase accountability for how district officials spend their time and allocate development projects.
- Increase citizen knowledge about and participation in the budgeting process for development projects.

We further discuss how each of these hypotheses are operationalized (Section 4) and tested (Section 5) below. We also discuss in the section on heterogeneous treatment effects that some of these relationships might be conditioned by the quality of the audit reports that the district received. In particular, strong or weak audit reports might condition the relationships we see between treatment condition and expected outcome.

3.0 RESEARCH DESIGN

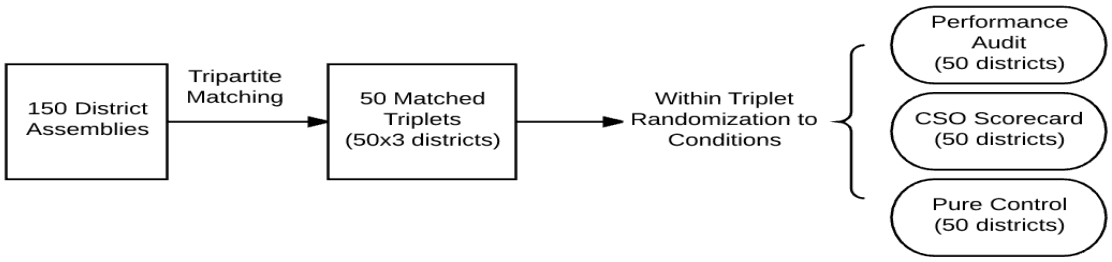
The study utilizes a randomized control trial design to test the impacts of each intervention of the program. To assess the individual effect of each intervention, the evaluation design is a three-arm district-clustered, triplet-blocked randomized control trial. Figure 4 illustrates the three treatment arms of the IE, consisting of the different interventions received (GAS Performance Audits with information campaign, CSO Social Audits with information campaign, or pure control). Districts were randomized into these three treatments across Ghana's 10 regions. A comparison of findings in GAS districts versus control and CSO versus control provides the average program impact on each of the components laid out in the hypotheses above.

3.1 MATCHING AND RANDOMIZATION PROCEDURE FOR GSAM

A quantitative matching algorithm was used for the random assignment of districts to GAS, CSO, and control groups. The goal of the matching and randomization procedure is to ensure that districts in the GAS, CSO, and control are as similar as possible. Absent initial matching, randomization could produce very different condition groups simply as a result of random chance. This potential problem is particularly stark with sample sizes below 300.[9] Increasing the similarity of districts prior to randomization offers several benefits. First, it decreases the risk that district level characteristics other than the randomly-assigned experimental conditions are responsible for any differences in outcomes between each of the three arms of the program. Differences between the treatment and control districts that are correlated with the program outcomes can lead to biased, incorrect inferences about program effects. Second, directly comparing districts that are most similar on pre-treatment district-level characteristics can decrease noise in the data and thereby improve statistical power and thus our ability to detect effects of the programming.

In order to ensure that the districts were as similar as possible, we conduct two procedures. First, we create triplicate matches of districts that are as similar as possible. Second, once we had all of the districts matched into similar triplicates, we randomly placed one district from each triplicate into the audit, CSO and control groups respectively. Figure 4 below provides a graphic representation of the matching and randomization procedure, which we further detail.

Figure 4: Evaluation Design



Matching required that we identify characteristics most likely to impact the quality of capital projects and political accountability across districts, i.e. the outcomes GSAM aims to improve. We selected six background

[9] See Bruhn and McKenzie (2008) and Greevy et al. (2004).

characteristics for the matching procedure that we expected to be strongly correlated with the outcomes of interest. In particular, we began the matching procedure by first using exact matching on geographic region (north vs. south) and whether a district is new (i.e. created from previous districts during Ghana's 2012 redistricting) or not. Following this exact matching, we selected a set of additional indicators most likely to predict the outcomes of interest. These included a district poverty index, the ethnic fragmentation of the district, the electoral competitiveness of the district, the district's most recent (2011) Functional and Organizational Assessment Tool (FOAT) score, and whether the district had previously experienced a GAS performance audit.[10] Having identified these key pre-treatment characteristics, we deployed a statistical algorithm to match the districts on the characteristics. A detailed explanation on our approach to triplicate matching is provided in our baseline report (Huntington, Schultz, and Wibbels 2015).

Once the district triplets were identified, the randomization procedure was straightforward. One district of each triplet was randomly placed in the GAS, CSO and pure control groups as determined by a random number generator. Within each triplet, the district with the highest number was assigned to the control group. The district with the middle number was assigned the CSO programming, and the district with the lowest number was assigned to the GAS audit programming.

The GSAM baseline report provides evidence that matching districts on pretreatment characteristics contributed to balance across control and treatment districts both in the characteristics of respondents and in survey responses (Huntington, Schultz, and Wibbels 2015). The midline data also shows good balance. In short, given data constraints the matching and randomization procedures did a good job of generating balance between treatment and control districts.

It is important to note that the study does not involve all of Ghana's districts, which numbered 216 as of 2017. Sixty-one of the 216 districts are classified as "metropolitan" or "municipal", meaning they are more urban, and these districts were excluded from the study. This was done to ensure the relative homogeneity and comparability of the three study arms. The programming was not aimed at these more urban districts. Of the remaining universe of 155 traditional districts, the 150 districts for GSAM were selected on the basis of the best matches as described above. As a result, our inferences concern the 150 districts in our sample rather than the broader population of all districts in Ghana. Although this may limit the generalizability of our results to other contexts, matching increases confidence in the unbiasedness of our estimates and should improve the precision of our estimates for a narrower population of districts.

3.2 SAMPLING

There were three stages of sampling: district, enumeration area (EA), and household. The district sampling process is described above while the enumeration area and household sampling is described below.

Household Enumeration Areas

At the baseline, the IE team selected the 150 districts and 50 matched triplets according to the criteria described in the matching procedure section above. At the second stage, two enumeration areas (the primary sampling units for the survey) were selected from each of the 150 districts into our sample. Using the 2010 Ghana Population and Housing Census Enumeration Areas (EAs) List as a sampling frame, the EAs in each of the

[10] These performance audits were done in ten MMDAs in conjunction with DANIDA.

selected districts were categorized into two strata (rural and urban). One rural EA and one urban EA were randomly selected from each district. The randomization structure was constrained to prevent the selection of neighboring EAs and those characterized by the Ghana Statistical Service (GSS) as "remote", i.e. unreachable by car or motorcycle. Overall, the proposed household sample size for the original RCT evaluation design was 3,000: 150 districts with 10 households sampled from each of the two EAs.

Deviations from planned intervention implementation (see section "Endline Challenges, Cautions and Implications for the Evaluation" below), required the IE team to alter the original IE plan. The GSAM audits—both GAS and CSO-led—and the resulting dissemination of citizen scorecards occurred in many villages that did not fall within the baseline survey enumeration areas, and many of the baseline EAs did not receive dissemination events. In other words, although audits took place in all districts according to plan, the dissemination of scorecard results did not cover many of the villages and towns in those districts.

If we rely only on the baseline enumeration area sampling for the endline impact evaluation, the resulting analysis will identify the effect of the intent to treat; however, the substantively more relevant quantity of interest in such an evaluation is the actual effect of the treatment itself. Given our knowledge of the actual implementation of the programming, restricting analysis only to the baseline enumeration areas could understate the effect of the program.

In order to address these challenges, the evaluation team added one additional enumeration area - comprised of 10 households - to the endline data collection in communities where the maximum suite of programming had taken place. In GAS audit districts, the additional EA was chosen from amongst those towns and villages where scorecard dissemination events had taken place. In the social audit districts, the additional EA was chosen from those villages where CARE and NGO partners had trained local citizens in social auditing and begun the process of diffusing scorecards. As a result of this additional EA in each of the treatment districts, the household survey sample increased from approximately 3,000 at baseline to approximately 4,000 respondents at endline. We discuss our approach to estimating program effects across the additional and original EAs in Section 6 below.

Administrative and Elected Official Surveys

The administrator and politician surveys aimed to conduct structured interviews with 750 DA administrators (5 per district) and 750 local political officials (5 per district) across the 150 study districts.[11] Due to variation in availability and accessibility of officials, the politician and administrative official surveys relied on systematically-obtained convenience samples of those most knowledgeable about district development projects and budgeting. In sampling officials for the political and administrative official surveys, enumerators were told to prioritize officials involved in the development and management of the district's development project budget. To facilitate consistency of sampling, the IE team provided the enumerators with a list of official roles and the prioritization of those officials. The list of Political Officials, from highest priority to lowest priority, included:

- District Chief Executive
- Presiding Member (PM) of the DA

[11] A limited midline survey was conducted in April and May of 2016 among administrators and political officials. The survey did not take place in the 50 social audit districts because the audits were still in progress, DA dissemination events had not occurred, and the citizen outreach component of the SA programming had not yet occurred. Thus, the midline report was based on a smaller survey of slightly less than 1,000 DA officials from the control and GAS districts. The midline survey instruments sought to assess whether overall levels of service delivery, corruption, governance effort, budget procedures, and other indicators of local government performance varied between control districts and the GAS-audited districts.

- Chair of the Works Sub-committee
- Chair Development Planning sub-committee
- Chair of the Finance Sub-committee
- Chair of the District Executive Committee (if different from DCE or PM of the DA)
- Chair of the Social Services Sub-committee

The Administrative Officials, from highest priority to lowest priority, included:

- MMDA Internal Auditor
- District Budget Analyst/Budget Officer Development Planning Coordinating Council
- District Coordinating Director (i.e. the head of the DA administration)
- Chair of District Tender Board (Note: only if it is not the DCE)
- District Finance Officer/ Comptroller
- District Planner/Director of the Physical Planning Department
- MMDA Planning Officer (if there is one)
- District Works Engineer
- Assistant District Engineer
- District Procurement Officer

This list of administrative officials is somewhat different from the list provided during the baseline survey. As the programming staff and the IE team learned more about district budgeting for development projects, we agreed to replace the Director of Social Welfare/Community Development and the Director of Education, Youth and Sports Development with the District Planner/Director of the Physical Planning Department and the District Procurement Officer. We also added the District Works Engineer and Assistant Engineer, who play an important oversight role during the construction of development projects. Enumerators and field supervisors were instructed to use these lists until they reached the target sample size of 10 officials for each district. Enumerators and field supervisors were also instructed to aim for balance between Political and Administrative officials in each district. District officials were contacted in advance of the survey team's arrival in the district in an effort to establish meetings with those higher on the prioritization list.

3.3 ENDLINE CHALLENGES, CAUTIONS, AND IMPLICATIONS FOR THE EVALUATION

Complications in program implementation had three important impacts on the IE team's capacity to test the effects of the interventions. First, the delay in the social audits created a mismatch in the timing between the two arms of programming. The GAS audits took place in the latter half of 2015 and the corresponding citizen outreach took place in the second quarter of 2016. The social audits, by contrast, were conducted throughout 2016 and the resulting scorecards and broader outreach efforts took place in the latter half of 2016. Endline data collection occurred in late-January to March of 2017, which means that we are testing the short-term impacts for the CSO treatment and the medium-term impacts for the GAS treatment.[12] This has implications for the theory of change. For example, if the theory of change is such that improved information and citizen engagement leads to pressure for public officials to better oversee development projects, which in turn results in improved projects, then it could be that insufficient time has passed to achieve this higher-level outcome. On the other hand, because endline data collection occurred so soon after the scorecards, the results of the intervention might be more salient in CSO districts than in GAS districts.

[12] As discussed below, data collection had to be conducted in early 2017 due to changes in district officials as a consequence of the 2016 national elections.

Second, while both interventions were initially conceptualized as district-wide interventions insofar as they involved dissemination activities across the entirety of the treated districts, the difficulty of reaching even a majority of citizens in programmed districts proved too challenging. Thus, there is considerable heterogeneity in the extent of programming within districts - some villages received extensive information campaigns on GAS audit results, while others did not. Likewise, some villages received extensive social audit training, other districts received the diffusion of social audit results, and yet others received neither.

The heterogeneity of the spatial concentration of information dissemination and social audit training has implications for our capacity to detect program effects. This is particularly likely to be the case at the household level. Most importantly, the randomly selected baseline EAs where household interviews took place were likely to have received relatively little direct information dissemination. The IE team addressed this challenge by adding an additional EA in the endline sampling. This EA was selected based on information provided by the implementers regarding the actual implementation of programming. More specifically, the additional endline EA was chosen from where we knew that programming took place. Third, two rounds of elections - one at the district level and the other at the national level - complicate the theory of change and our capacity to identify program effects. The first elections - a set of DA elections, took place in October of 2015 after having been originally scheduled for the previous year. More than half of political officials in the baseline sample of GAS audit districts were newly elected in October 2015. These newly elected officials were not present for the GAS audits or the exit meetings regarding the audits. Though these elected officials might have experienced subsequent citizen mobilizations as a result of GSAM and did attend DA dissemination events, their more limited exposure to programming might attenuate program effectiveness.

Fourth, the delay in programming resulted in it running up against the second set of elections - the December 2016 presidential election. The 2016 presidential election resulted in the loss of the incumbent president John Dramani Mahama to challenger Nana Akufo-Addo. The presidential election consumed much of the energy and attention of the Ghanaian government. The passage of DA budgets was delayed, which complicated the timing whereby citizen pressures might be brought to bear on the development of DA Annual Action Plans. Moreover, the resulting change in the party of the president has important implications for the composition of district governments, since the president appoints the key DA political official - the District Chief Executive (DCE) - and a third of the members of the District Assembly.

Given that newly appointed DCEs and DA members would have no history with GSAM, this threatened the IE team's capacity to measure the impact of GSAM on political officials. Thus, the IE and field teams had to rush to get the endline into the field, and the Institute of Statistical, Social and Economic Research (ISSER, the local data collection partner, had to hire a particularly large survey team of 61 enumerators, in order to reach outgoing DA political officials before the turnover. These efforts were mostly successful, although the pending turnover did make it more difficult to gain interviews with the requisite officials and the household survey failed to reach GSAM communities in 8 of the GAS districts.[13] In addition, significant turnover in DA officials was in the process of taking place throughout the duration of endline data collection and might have impacted the willingness and responses of outgoing officials to respond to questions about DA budgeting and development projects and also affected the willingness of incumbent officials to provide responses on controversial district issues most relevant to the intervention's outcomes - such as issues in district budgeting and contracting.

[13] Household interviews were conducted in baseline EAs in these eight districts, but a third, GSAM-intensive EA was not enumerated. These include Kadjebi, Krachi Nkhurmuru, Krachi West, Amansie West, Asante Akim South, Kumawu.

Fifth, as noted in the explanation of the interventions, the efforts to dissemination GAS audit scorecards to citizens was more extensive than in the CSO communities.

Sixth, the original GSAM design was only moderately powered (see baseline report). Despite sampling large numbers of political officials, administrative officials, and households during data collection, the fact that assignment to experimental conditions occurs at the level of the district limits the precision of our estimates. The challenge of detecting moderately sized and small treatment effects becomes more difficult when we attempt to look for heterogeneous impacts. For example, it is likely that citizens in districts with positive audit results will have a different view of their local government than citizens in districts with negative audit results, but our ability to test for this heterogeneous impact is limited by the relatively small number of districts.

4.0 DATA

Baseline, midline and endline data collection include: (1) a household survey, (2) an administrator survey, (3) a politician survey, and (4) administrative data, including Annual Progress Reports issued by district governments and the GAS's 2016 report on financial irregularities in the districts.

4.1 QUANTITATIVE DATA SOURCES

For baseline and endline data collection, the household, politician and administrator surveys are structured quantitative instruments. The household survey collects data on a number of measures including livelihoods, satisfaction with local governance and service delivery, quality of district capital projects, responsiveness of district officials and citizen satisfaction and engagement with district-level democratic procedures, in addition to basic social, economic, and demographic characteristics. The politician and administrator surveys seek to assess procedures governing development projects, corruption, governance effort, and adherence to budget procedures.

The household, politician and administrator surveys were collected using electronic data capture. All data was collected by the ISSER, a research center housed at the University of Ghana, in the local languages of each region. At endline, the evaluation team led a one-week training of field staff (many of whom had worked on the baseline and midline) at endline. Endline data was collected during a two-month period from late January – March 2017.[14] Endline data collection was subject to a series of quality checks. Given the rush to complete data collection in a timely manner, surveys themselves were not audited by field managers as they were at baseline. A series of regular checks conducted by the evaluation team screened the data for irregularities as data collection progressed.

The study made every effort to re-interview baseline household survey respondents at endline for the original 300 baseline EAs. The survey firm was provided with baseline respondent names, addresses, phone numbers, and landmarks to help identify the correct household panel respondent. (Personal data was stored on a protected data network.) Despite this effort, the evaluation team was only able to verify 1,605 household matches to the 3,000 baseline respondents, including 1,264 exact panel respondent matches. This smaller than expected panel sample may be the result of a high level of survey attrition and/or data quality issues in the survey firm's household tracking information, which hindered its ability to verify matches.

Lastly, we move beyond opinion data by analyzing two types of objective outcomes at the district level. First, the IE team digitized DA-generated Annual Progress Reports for 2016. These reports provide information on the location, type, funding source, and level of completion of district development projects. DAs nearly always submit them in a timely manner because they are a prerequisite for receiving central transfers in the current fiscal year. Since one of the most important sources of waste in DA budgets is abandoned projects (Williams 2016), we calculate the share of projects that are incomplete for 2016, the most recent year for which data is available. If GSAM improves administrative and public oversight of projects, it should be reflected in a lower rate of project non-completion. Second, the IE team digitized the Ghana Audit Service's annual report on district management of public funds in 2016. Though this financial audit does not speak directly to district management

[14] A limited midline sample of administrators and political officials was collected in the 50 GAS audit districts from March – April 2016. These data are not used for the endline analysis given the limitations in geographic scope.

of development projects, it does assess the extent of irregularities in DA finances. To the extent expenditures on development projects represent the lion's share of DA spending, GSAM could have had an impact on the extent of financial irregularities.

As is common, missingness exists in the survey data that underpins our outcome measures. We drop from the analysis questions for which more than 30 percent of responses were missing. When missingness was lower, we impute missing values using the method proposed by Honaker and King (2010); we use religion, level of education, ethnicity, an asset index, region, and questions bearing on how often respondents follow news and have difficulty getting enough food and clean water, NDC vote share in 2012, as well as the outcome measures as predictors.

4.2 ETHICAL CONSIDERATIONS

Participation in the study was voluntary, and all respondents were required to give their informed consent at the beginning of the survey process. The GSAM IE team received Institutional Review Board approval for the survey instruments from SI on June 20, 2014 and again on December 29, 2016; minor modifications were made to questions for the endline surveys. Informed consent was received from each participant after reading a statement about the purpose of the research, the content of the survey, any risks or benefits, and the time commitment. Participants were assured their participation was voluntary and could be withdrawn at any point and that their answers would be kept confidential.

As described above, quantitative data was collected through the ODK platform on Android tablets. Tablets were password protected, and data was uploaded to an encrypted server when network connectivity was available. Data is stored on password encrypted computers, with PII removed. Identified data will be stored on Duke University's Protected Data Network.

5.0 INDICATORS

Per the hypotheses outlined in Section 2.4 above, this study examines the effect of GSAM on the three sets of primary outcome measures and a fourth focused on awareness of the GSAM program. In addition to these primary outcome indicators, we construct a series of secondary measures that augment our understanding of the program's effects. In several cases, we rely on multiple related survey items bearing on a core concept to reduce the noise of measurement. In such cases, we follow the standard approach of calculating z-scores as summary measures (Kling, Liebman and Katz 2007).[15] Below, we restate the relevant hypotheses and define the relevant measures. Table 2 (household survey) and Table 3 (administrator and politician survey) in Annex II provide a summary of the indicators and details on how the IE team created indices for those models where we rely on multiple measures of the same construct. In a number of cases, we focus specifically on primary ("basic") schools because these represent far and away the most common DA-built development projects.

OUTCOME FAMILY 1: CAPITAL PROJECTS AND SERVICE DELIVERY

Outcome Family 1 is focused primarily on a series of variables pertaining to the hypothesis that GSAM will improve how citizens perceive the quality of district capital projects, as well as health and education service delivery. Indicators from Outcome Family 1 are constructed out of the household survey and test the following: The interventions will increase ***citizen*** benefits from development projects and the corresponding services, reduce perceived obstacles to completing projects, and increase citizen satisfaction with services. All indicators are explained in greater detail in **Error! Reference source not found.** and **Error! Reference source not found.** in Annex II.

Primary Indicators (citizen survey):

- **Benefits from development projects**: Index of family and community benefits from recent development projects.
- **Satisfaction with school facilities**: Ordinal measure of satisfaction with physical condition of school facilities.
- **Satisfaction with building projects**: Index of citizen satisfaction with DA building of: a) roads; b) health facilities; and c) schools.
- **Perception of delays**: Ordinal measure of the perception of the frequency with which work delays happen on development projects.
- **Satisfaction with services**: Index of respondent satisfaction with basic education, health services and public toilets.

Within **district administrations**, i.e., among administrators and political officials, the interventions is hypothesized to improve district planning, contracting, and oversight of development projects.

Primary Indicators (administrator and politician survey):

[15] These are calculated by: a) recoding all variables such that better outcomes have higher scores; b) standardizing each recoded variable by subtracting the control group mean from the district-level average, and dividing by the control group standard deviation; and c) averaging across the standardized variables.

- **Internal DA oversight:** Index of the strength of internal institutional oversight of projects as measured by the existence and influence of a DA Audit Report Implementation Committee.
- **Frequency of project abandonment**: Continuous measure of the share of development project that are reported abandoned by administrators and political officials.
- **Project completion rate:** Continuous measure of the percent of development projects completed in 2016, as collected from *Annual Progress Reports*.
- **Project outcomes**: Index of the quality of project implementation over the past year as measured by building delays, abandonments despite payment, and on-time completions.

OUTCOME FAMILY 2: DISTRICT GOVERNANCE QUALITY

Outcome Family 2 tests the hypothesis that GSAM improves district governance quality. This hypothesis is assessed through variables measuring the quality of district governance, with special reference to district capital projects. Indicators from Outcome Family 2 test whether the interventions reduce the incidence of corruption, increase perceived government transparency**,** and improve citizen's perceptions of the relative performance of district governance.

Primary Indicators (Citizen survey):

- **Obstacles to development**: Index of perceived governance obstacles to development as measured by citizen perceptions that: a) lack of citizen participation; and b) corruption among district officials are obstacles to building better development projects.
- **Relative performance**: Index of relative governance performance as measured by citizen perceptions of DA: a) corruption; and b) inefficient spending, relative to other Das.
- **Perception of personal enrichment**: Index of perceived personal enrichment by DA officials, including: a) the DCE; b) elected DA member; and c) DA professional staff.
- **Perception of partisan bias:** Survey experimental evidence on the extent to which DA officials perceive members of their party as providing privileged access to public resources to other party members.

Primary Indicators (administrator and politician surveys)

We also test our hypothesis on data collected from district administrators and politicians:

- **Corruption impeding development**: Index of the extent to which: a) corruption is a problem in the preparation of Annual Action Plans; b) corruption is a problem in the execution of Annual Action Plans; c) corruption in the DA is an obstacle to building development projects that serve the needs of citizens; d) corruption among contractors is an obstacle to building development projects that serve the needs of citizens.
- **Project transparency**: Index of transparency in: a) the planning and b) the contracting of district development project.
- **Relative performance** Index of relative governance performance as measured by perceptions of: a) corruption and b) inefficient spending, relative to other DAs.

OUTCOME FAMILY 3: DISTRICT GOVERNMENT ACCOUNTABILITY AND CITIZEN ENGAGEMENT

Outcome Family 3 tests the hypothesis that GSAM improves government accountability and citizen engagement. Outcomes are evaluated through variables test whether interventions will improve the responsiveness of DAs to community needs when it comes to budgeting and locating development projects.

Primary Indicators (Administrator and Politician surveys):

- **Need as criteria for project location:** Indicator variable of the importance of targeting communities with the most needs in the allocation of development projects.
- **Perception that public participation will improve development**: Ordinal measure of the extent to which the lack of citizen engagement represents an obstacle to better development projects.
- **Strength of public complaints committee**: Index of the existence and activism of the DA Public Complaints Committee.

Primary Indicators (Citizen survey)

- **Project location**: Ordinal measure of perceived importance of community needs in where the DA locates projects.
- **Opportunities for public input**: Index of whether community was consulted on the largest recent project in the area, whether the DA or town council hold public meetings to establish development priorities.

The interventions will increase accountability for how ***district officials*** spend their time and allocate development projects.

Primary Indicators (Administrator and Politician surveys):

- **Time spent on the public**: A continuous measure of the share of work time that administrators and politicians spend responding to citizen concerns and/or working with community/civil society groups.

The interventions will increase citizen knowledge about and participation in the budgeting process for development projects

Primary Indicators (citizen survey):

- **Participation in project priority meetings**: Whether respondents have participated in community meetings aimed at developing priorities for development projects.

AWARENESS INDICATORS

In addition to these measures to test our hypotheses, we also assess the extent to which citizens and DA officials in programming committees are aware of GSAM, whether it impacted their engagement with the DA, and their overall assessment of it. Analysis of these questions serves two purposes: 1) if GSAM is to have any direct effect, it must be via some awareness of the programming; and 2) as discussed in Section 2 above, there is heterogeneity in the extent of programming across communities within programmed districts and potential spillovers between them. The measures include:

Citizen Survey:

- Whether or not the respondent has heard of GSAM
- Whether or not the respondent has attended a GSAM meeting
- Whether or not the respondent has contacted a DA official due to GSAM
- Whether or not and how the respondent's involvement in DA politics has changed due to GSAM
- The respondent's overall assessment of GSAM
- Whether or not respondents think GSAM should continue

Administrator and Politician Surveys:

- Whether or not the respondent has heard of GSAM
- Whether or not citizens have mentioned GSAM to the respondent on others in their office
- Whether or not other members of the district government have mentioned GSAM to the respondent on others in their office

- Whether or not and how the district government has taken any action in response to GSAM
- Whether or not and how citizen engagement with district government has changed as a result of GSAM

Indicator Table 2 and Table 3 in Annex II provide further details on how each of these measures are produced. Additionally, Annex II provides details on a host of secondary outcomes that the IE team collected data on. We test for the impact of GSAM on these secondary outcomes and report on them, albeit in less detail, below.

TREATMENT INDICATORS

We test for the effect of two different treatments and one form of heterogeneous treatment effects. They are measured by the following indicators:

- **GAS audit indicator**: An indicator variable that equals 1 if a district was assigned to receive the GAS audit intervention, 0 if a district was not assigned to receive the GAS audit intervention.
- **CSO SA indicator**: An indicator variable that equals 1 if a district was assigned to receive the Social Audit intervention, 0 if a district was not assigned to receive the Social Audit intervention.
- **Development project EA**: An indicator variable that equals 1 for development project enumeration areas in treatment districts, which indicates high-intensity treatment communities, i.e. those communities in GSAM districts where we know the programming occurred. This is only produced for household analyses.

Interested readers can find descriptive statistics for all of these variables in the online appendix.

6.0 ESTIMATION STRATEGY

PRIMARY ANALYSIS I: MAIN EFFECTS FOR CROSS-SECTIONAL ENDLINE SAMPLE

The IE is designed to rigorously assess the direct impacts of the GSAM interventions on the three outcome families described above and awareness of GSAM. The analysis tests the impact of GSAM on the outcomes described in Section 5 ("Indicators") at the household and community level. Our estimation will compare the following two groups:

Group 1) Each treatment arm (GAS audits, social audits) to the pure control.

Group 2) GAS audits to the CSO-led social audit campaign.

The primary analysis relies exclusively on endline data. We do this for three reasons. First, we had good balance between the three arms of the study at baseline, indicating that observed differences at endline would be due to the intervention rather than baseline differences. Second, we only have endline data for the enumeration areas that were added based on treatment location. Third, due to changes in the program design since baseline, some updates were made to the endline instruments to better capture program effects. This restricts the number of indicators that can be used for panel analysis because of differences across the baseline and endline questionnaires.

In this primary analysis we present results of the most straightforward approach to data analysis, which reports the intention-to-treat effect (ITT) as the difference in mean outcomes between treatment and control districts without "control variables". As discussed in Section 2.3 above, there was considerable heterogeneity in the extent of treatment within districts. While both interventions were initially conceptualized as district-wide, diffusing information across all, or even most, district citizens proved too challenging. This heterogeneity is most marked in the extent of information dissemination across communities within GAS and SA districts, where many communities received no dissemination events. The use of radio jingles and billboards, moreover, makes it impossible for the IE team to know exactly how much programming each community had experienced. As described in Section 3.4 above, the IE team addressed this ambiguity by supplementing the baseline household enumeration areas with an additional one where we did know that programming had taken place. These EAs were chosen from a list of communities that implementers targeted with a full range of community dissemination. We use Ordinary Least Squares (OLS) to estimate the effect of each treatment arm of interest using the following specification:

$$[1] \qquad Y_{ijd} = \beta_0 + \beta_1 T_j^{GAS} + \beta_2 T_j^{SA} + \beta_3 E_j^{SA} + \beta_4 E_j^{GAS} + \varphi_d + u_{ij}$$

where Y_{ijd} is the outcome measure of household i in district j in triplicate d. As described in Section 5, most outcome indicators are constructed from survey data and can be continuous, dichotomous, or ordinal. We also include matched triplicate fixed effects ϕ_d per Bruhn and McKenzie (2009) to account for design effects.[16] T_j is the treatment dummy for each of the two treatment arms of interest, and E_j is a dummy variable for the third

[16] See Bruhn and McKenzie (2009), who show that failure to control for the method of randomization typically produces overly conservative standard errors.

EAs in SA and GAS districts.[17] $\mathcal{u}_{ij}$ are robust standard errors clustered at the district level using Huber-White sandwiched standard errors (Lin et al., 2013). We rely on OLS (rather than a combination of linear and non-linear models) across all these variable types in light of evidence on the robustness of OLS in experimental settings. [18]

In the Appendix we present a host of alternative specifications to test the robustness of our findings. The key alternative specifications report:

- Estimates of the conditional average treatment effect (CATE) per equation 1, estimated separately above and below the mean audit score. One potential source of heterogeneity in the impact of GSAM is the quality of the district audit scores that were reported to citizens and district officials via scorecards. The interventions might improve citizen's perceptions of district governments, reduce perceived corruption, and increase perceived transparency when district scores on GAS or CSO audits are positive. By contrast, the interventions are likely to decrease the outcomes above when district scores on GAS or social audits are negative. We test for these potential heterogeneities with a dummy variable, which takes on a value of one when district scores are above the median and zero otherwise. Since audit scores are not available for control districts, we cannot interact them with treatment status. Instead we estimate the CATE above and below the mean audit score.
- The ITT without separately modeling the additional GSAM EAs in GAS and CSO districts.
- Models with a set of "control" variables. At the household level, these controls include a household asset index, calculated as the first component in a principal components analysis conducted on a series of asset questions and key respondent characteristics, including gender, age and ethnicity. At the district-level, we control for NDC presidential vote share in 2012 since politicized targeting of development resources could impact our outcomes of interest. In the analysis of the Administrator and Politician data, we introduce a measure for ethnicity and the length of time respondents have been in office. This latter addresses the potential for: a) newly rotated administrators into treatment districts who have experienced little programming; or b) for newly rotated administrators into control districts who were previously in treatment districts, who might have extensive experience with GSAM.

We note in the body of the main report when these results deviate from (or shed light on) those reported in the tables and figures below. These results are reported in an online appendix.

SECONDARY ANALYSIS: MAIN EFFECTS FOR PANEL DATA

The secondary analysis for the evaluation focuses on a panel analysis of household data, including both baseline and endline data.

Equations [3a] below is applied to the analysis of the panel data:

$$[3a] \qquad Y_{ijd} = \beta_0 + \beta_1 T_{ij}^{GAS} + \beta_2\ T_{ij}^{CSO} + Y_{ij}^{0}\ + \phi_d + \mathcal{u}_{ij}$$

where Y_{ijd} is the outcome measure of household i in district j. T_{ij} is the treatment dummy for each of the two treatment groups of interest. Y_{ij}^{0} is the baseline vectors for the outcome measure, ϕ_d is the matched triplicate

[17] A dummy variable means that the variable is coded one if it has the attribute (e.g., GAS treatment) or zero if it does not.
[18] See Judkins and Porter (2016).

fixed effects, and u_{ij} are robust standard errors clustered at the district level, using Huber-White sandwiched standard errors (Lin et al., 2013).

Due to panel attrition between baseline and endline, the household analysis will explore outcomes across two subsets of the data:

- A respondent panel, including only exact matched respondents between the two survey waves;
- A combined household and respondent panel, including all observations for the two subsets noted above.

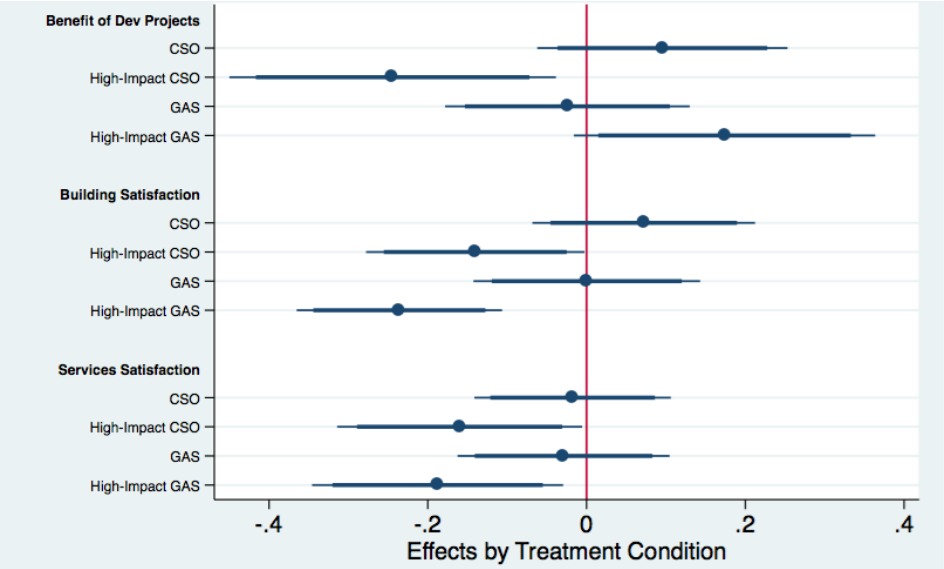
7.0 RESULTS

In presenting results, we provide graphic presentations of the estimated program effects for GAS and CSO treatment districts, as well as separate estimates for the "high-intensity" GSAM communities. These results reflect equation 1 above. We organize our presentation of results by the data source, (a) household results, (b) administrator results, (c) politician results, and district outcomes, and by the three outcome families discussed in Section 5 above, (a) development projects and service delivery, (b) district governance quality, and (c) accountability and citizen engagement. We also examine program effects on awareness of GSAM. The figures present point estimates and 90 and 95% confidence intervals; if confidence intervals cross zero, the reader should interpret the results as null. In many cases, the outcome variables are measured as scales and the point estimates represent effects in terms of standard deviations. In graphing results, we focus on those cases where we uncover significant program effects. We discuss nulls in the text and refer the read to the full presentation of tabular results in Appendices 3 (full household results), 4 (full administrator results) and 5 (full politician results) of the online appendix.

7.1 PRIMARY ANALYSIS: HOUSEHOLD RESULTS

Outcome Family 1: Development Projects and Service Delivery: Outcome Family 1 focuses on development projects and service delivery, Figure 5 below shows that GSAM's effects are only measurable in high-intensity communities. More detailed regression tables are presented in Annex IV. Across other CSO and GAS districts, the program has no impact on the perceived benefits of development projects, satisfaction with school facilities, the extent of project oversight by the DA, satisfaction with DA construction of roads, schools and health facilities ("Building Satisfaction" in Figure 5), or overall satisfaction with services. Focusing on high-intensity GAS communities, the program is associated with a .18 standard deviation *increase* in the perceived benefits of local development projects (significant at .10). At the same time, the program is associated with a .24 standard deviation *decrease* in satisfaction with DA construction, and a .18 standard deviation d*ecrease* in satisfaction with DA-provided services. These results on building and service satisfaction are echoed in high-intensity CSO communities, where GSAM *decreases* satisfaction by .16 and .19 standard deviations, respectively. Interestingly, GSAM *decreases* perceived benefits from development projects in high-intensity CSO communities; this is the one case across all of our results when CSO and GAS programming seem to have divergent effects. We suspect this might result from the intensive training on project oversight and quality that local citizens received in high intensity CSO communities; this training might have made citizens more critical of project quality than they would have otherwise.

Figure 5: Selected Development Project and Service Delivery Effects among Citizens



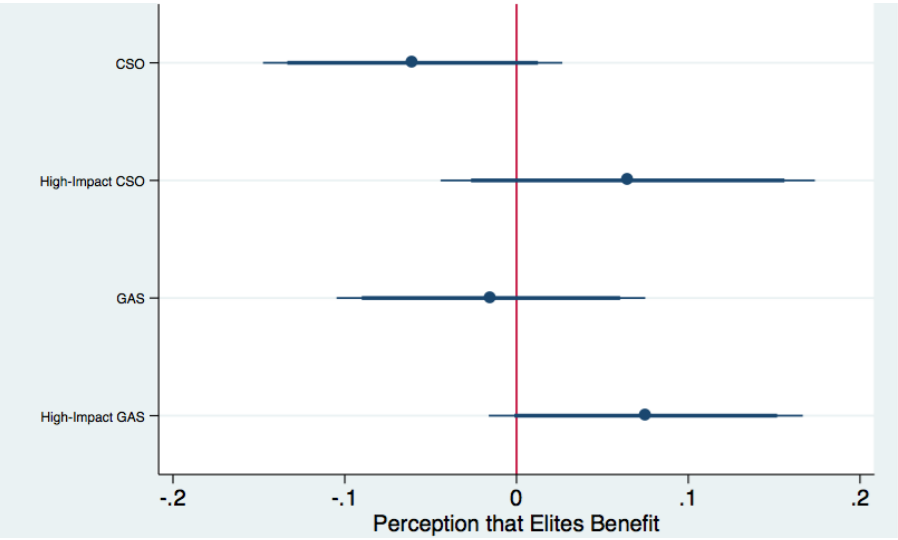
All told, these results suggest that in those communities where programming has been most extensive, GSAM has a mostly negative impact on satisfaction with DA project building and services. This could be a positive development from the point of view of accountability, if citizen dissatisfaction is a result of poor performance by district governments and a precursor to mobilizing for change. Further analysis suggests that dissatisfaction is driven by the poor performance identified in many audit reports. We compare results across districts with above average and with below average scorecards and find that the negative effect of GAS programming on service satisfaction is nearly twice as large (-0.25 standard deviations) for high-intensity GAS communities that received low audit scores. We find similar results in CSO communities with below average audit scores.[19] To the extent this correct attribution of responsibility for negative audit results—itself an important development from the point of view of accountability—contributes to additional engagement with DA officials, GSAM offers the potential for sustainable change.

Outcome Family 2: Governance Quality: We find that neither GSAM intervention had a statistically significant effect on governance quality indicators, including perceptions of corruption and a lack of citizen participation as obstacles to development, knowledge of district performance relative to other districts, or perceptions of personal enrichment by DA officials. Estimated effects on this last indicator are presented in Figure 6. We also do not find an effect in high-intensity EAs.[20] It is also the case that these main effects seem not to be driven by heterogeneous treatment effects bearing on audit quality.

[19] For building satisfaction, the magnitude of the GSAM effect is estimated to be about the same for high- and low-ranked districts, but there is greater variation in the high-ranked districts. More consistently negative assessments in low-ranked districts means tighter confidence intervals and more significant effects for high-impact CSO and GSAM programming in these lower performing areas.

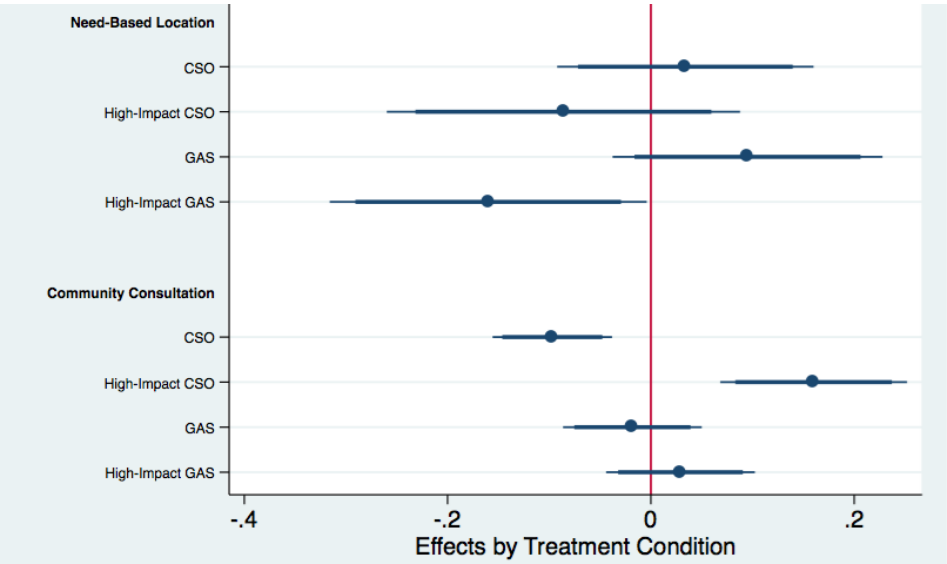
[20] The appendix shows that the GAS treatment does have a significant, positive effect on citizen perceptions of financial enrichment by DA officials in high intensity GSAM communities, but only when individual-level control variables are included in the model.

Figure 6: Selected Governance Quality Effects among Citizens



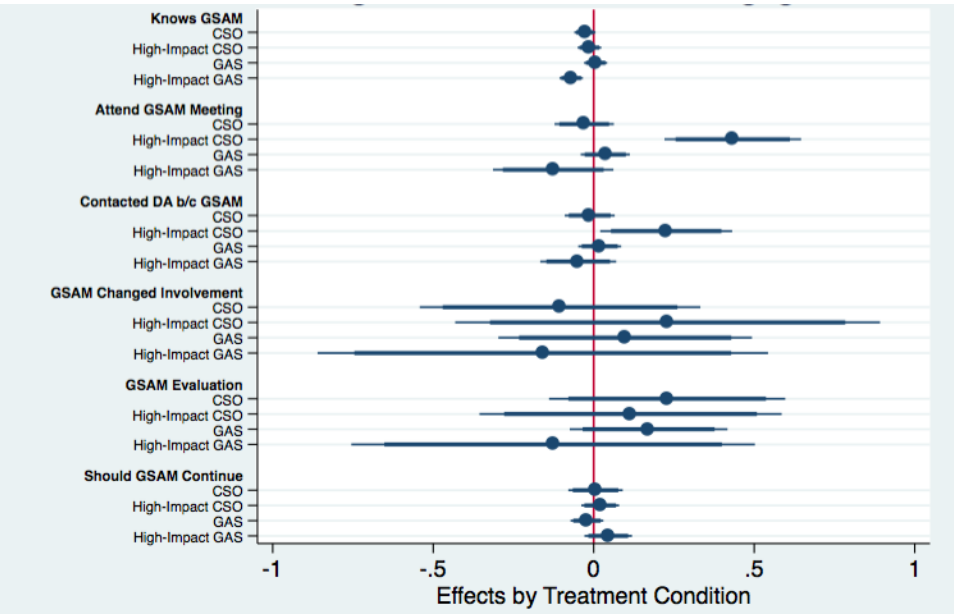
Outcome Family 3: Accountability: Figure 7 below presents results for *Outcome* 3, which focuses on citizen engagement and the perceived responsiveness of DAs to citizens. Some of these results are null—GSAM districts and communities do not perceive more opportunities for citizen input, and citizens are not more likely to attend meetings aimed at planning development projects; however, there are two results that are suggestive of improved accountability. First, in high-intensity GAS communities, there is a significant, .18 standard deviation *reduction* in the perceived role of community need in how DAs allocate projects. Again, however, this negative effect reflects only those communities that received negative audit scorecards; when scorecards are positive, the effects are null. In high-intensity GAS communities, the effect in low-scoring districts is three times more negative than in high-scoring districts, and it is only in these low-ranked districts where the effect is statistically significant. Put simply, poor audit scores translate into worse perceptions of how the DA responds to community needs. The other positive outcome from the point of view of accountability is that high-intensity CSO communities are .18 standard deviations more likely to report that their communities were consulted in the development of the largest, recent development project.

Figure 7: Selected Accountability Effects among Citizens



GSAM Awareness: Finally, Figure 8 below presents results for GSAM awareness. One important finding is that despite concerted efforts to diffuse information about the program in GSAM districts, those efforts have not reached a very large share of citizens; this may help explain why the findings discussed above are concentrated in high-intensity GSAM communities. Only 8.6% of all household respondents have heard of GSAM ("Knows GSAM" in Figure 8), and that share does *not* increase in either CSO or GAS districts; indeed, when we restrict our view to "high-intensity" GAS communities, the numbers are event slightly decreased. Across all respondents in CSO and GAS districts, Figure 8 shows no impact on having attended a GSAM meeting, contacting the DA as a result of GSAM, or changing engagement due to GSAM. The results on GSAM engagement are more encouraging when we focus on high-intensity CSO communities; among those who know of GSAM, respondents are 43% more likely to have taken part in a GSAM community meeting and 23% more likely to have contacted their DA as a result of the program.

Figure 8: GSAM Awareness Effects among Citizens



7.2 SECONDARY ANALYSIS: HOUSEHOLDS

Turning to the secondary analysis described in the pre-analysis plan, which uses panel data for both baseline and endline, we similarly find variation in whether outcomes are significantly correlated with GSAM. Where there are significant findings, they are mostly encouraging. Perception of government obstacles to development decreases, as does the belief that party or ethnic ties are needed for development projects to succeed. Politicians are held more responsible for services like schools, and politicians promise more community-oriented versus private benefits when campaigning for votes.

Focusing on *Outcome Family* 1, the **panel analysis** does not find any evidence of a significant treatment effect on primary or secondary service delivery outcomes for GAS or CSO areas. In the **endline survey** analysis, respondents in CSO districts, regardless of GSAM intensity, report that incomplete projects are slightly less important in their districts, compared to respondents reports in control districts. The difference is small but significant: two percentage points lower for CSO versus control districts. Differences for this outcome are not significant for GAS districts compared to controls.

For high-intensity GAS areas only, GSAM is associated with a 0.2 standard deviation *larger* walking distance respondents reported to three different services: nearest health facility, public toilet, and basic school. Given the timeframe of GSAM implementation and its evaluation, it is unlikely that this result reflects an actual change in the availability of services links to GSAM. Rather, consistent with our qualitative knowledge of programming within districts, it is possible that high-intensity GSAM areas were selected by implementing partners due to a lower baseline level of services and/or greater need for projects.

For *Outcome Family* 2 (governance quality), the **panel analysis** finds evidence of fewer governance obstacles to development in CSO districts, which represents a primary indicator for the evaluation, as well as greater satisfaction with district officials in GAS districts. Secondary results for the **endline analysis** show that in CSO districts, particularly high-intensity communities, respondents say that politicians promise more community development projects than private gifts in hopes of winning people's vote.[21] This may be due to increased visibility and importance of development projects provided by CSO auditing.

In GAS districts, the **panel *and* endline** analyses indicate that GSAM decreases the perception among citizens that their current DCE would likely help members of their own party *or* ethnic group at the expense of the public in the district. The size of this effect at endline is approximately -0.1 standard deviation for both measures, and, for ethnic favoritism specifically, the effect appears to be especially significant for lower-intensity GSAM areas. The reduction in ethnic favoritism is also evident in CSO districts for the panel findings.

Taken together, these results suggest that GSAM has some positive effects on the perception that community development projects are politically important and that projects can be implemented without regard to party or ethnic affiliation. However, it also seems that high-intensity programming may be associated – at least in CSO districts – with expectations that politicians will pocket a larger portion of financial windfalls earmarked for community development. High-intensity CSO areas expect politicians to pocket 5% more of the money compared to control districts.

[21] Respondents were asked how often politicians have promised benefits from 1) community development projects and 2) private money or gifts. The difference between the two responses reflects the relative magnitude of community versus clientelistic promises (although it does not tell us in absolute terms to what extent promises of any kind are being made in exchange for votes in a given district).

For *Outcome Family 3* (accountability), the **panel analysis** provides limited evidence that GSAM has promoted participatory government and improved DA responsiveness to citizen needs. In GAS districts, households are more likely to perceive that development project location is based on community needs. In line with the primary endline analysis, there also is panel evidence of greater citizen engagement in district governance for CSO areas.

In the **endline analysis**, for both CSO and GAS districts, respondents say that politicians are more responsible for the quality of projects such as local school buildings. The effect size for both audit groups is about 0.15 standard deviations larger than for the control group. In the CSO condition, the effect is observed regardless of programming intensity in the area, while in the GAS condition, the effect is larger for low-intensity areas and null for high-intensity areas.

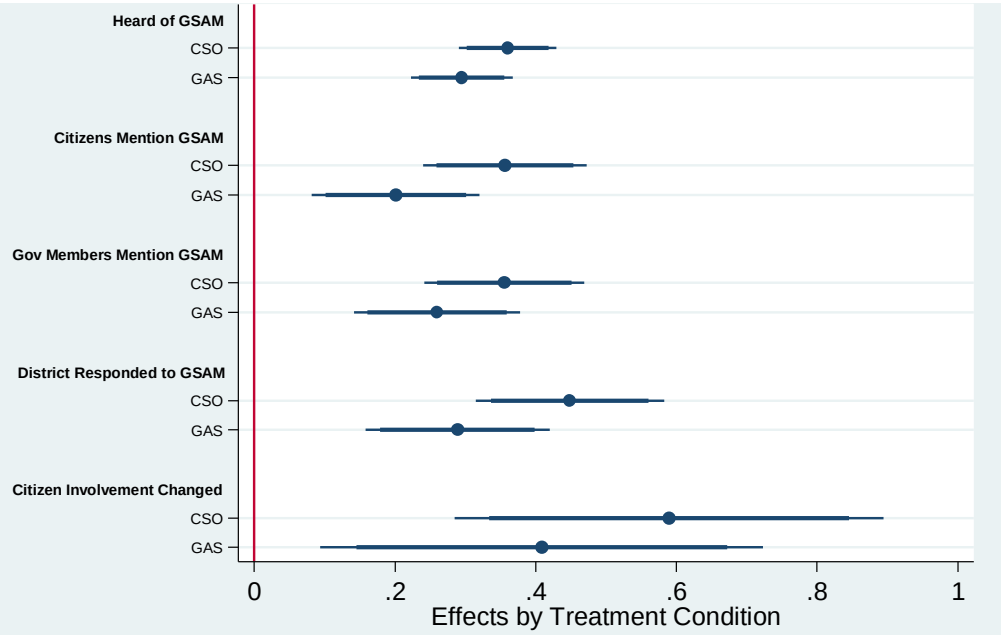
Furthermore, in both CSO and GAS areas, respondents are more likely to say that they would first turn to an *elected official* for help with their priority development project. In the GAS condition, the effect is observed regardless of programming intensity in the area, while in the CSO condition, the effect is larger for low-intensity areas and null for high-intensity areas.

In general, the IE team is reassured to see that the primary analysis (endline only) and secondary analysis (baseline and endline) provide fairly consistent findings.

7.3 ADMINISTRATOR RESULTS

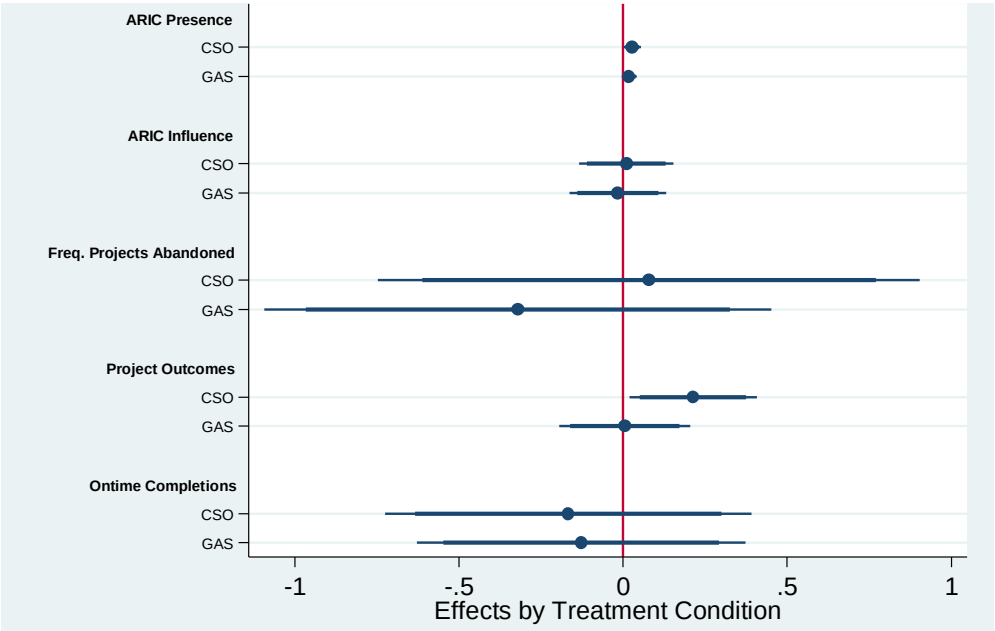
GSAM Awareness: Administrators in both CSO and GAS districts show a high rate of familiarity with GSAM (Figure 9 below). 75% of the administrators surveyed in GAS districts and nearly 81% of administrators in CSO districts reported that they had heard of GSAM, compared to 45% of administrators in control districts. For the subset of administrators familiar with GSAM, we see significant, positive results on all of our indicators of GSAM engagement. Administrators in GSAM districts are more likely to say that citizens have mentioned GSAM to them (20-40 percentage points higher); more likely to say that other officials have mentioned GSAM to them (25-35 percentage points higher); more likely to report taking action in response to GSAM (30-45 percentage points higher); and more likely to perceive higher citizen engagement due to GSAM. For this last measure, effects were especially large: citizen engagement due to GSAM was reported as being 40 percentage points higher in GAS districts and a full 60 percentage points higher in CSO districts. As clearly illustrated in Figure 9, across all indicators in this category, effects are slightly larger for CSO districts than in GAS districts. In the cases of citizens mentioning GSAM and districts responding to GSAM, these differences in effect sizes related to CSO and GAS programming are statistically significant.

Figure 9: GSAM Awareness and Engagement Effects among Administrators



Outcome Family 1: Capital Projects and Service Delivery: Despite strong familiarity with GSAM in target districts, as well as perceived improvements in discussion of GSAM, district responsiveness, and citizen involvement due to the programming, results are more limited for how GSAM affects program outcomes. Figure 10 below shows the results for Administrators in CSO districts are more likely to report having an Audit Report Implementation Committee (ARIC) in their district, though note that the vast majority of respondents confirm the existence of ARICs; effects are not significant for GAS districts. Among those who do report having an ARIC in the district, GSAM is not associated with any difference in ARIC influence.

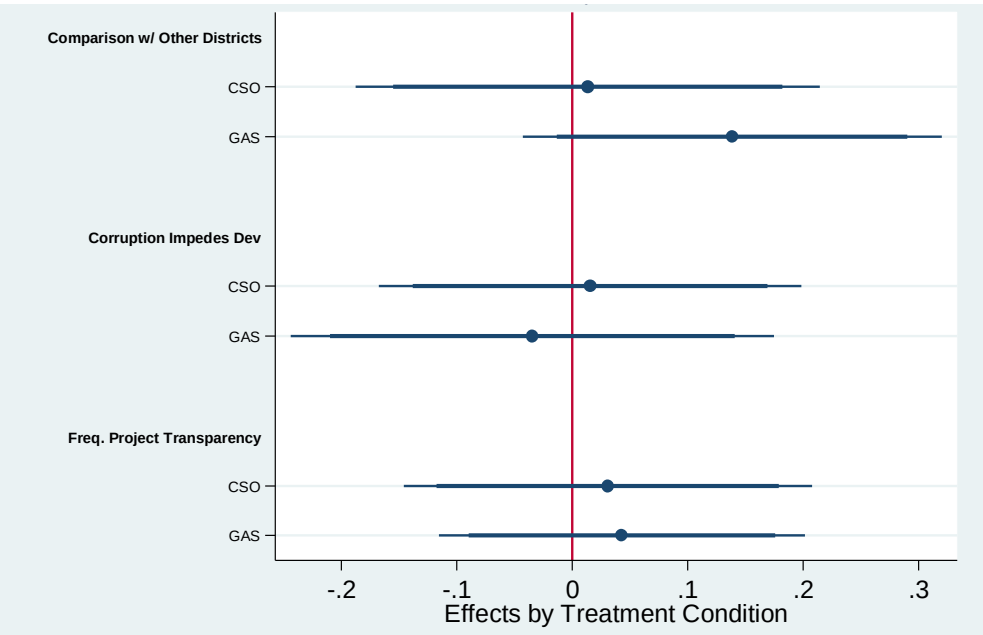
Figure 10: Development Projects and Service Delivery Effects among Administrators



CSO programming is associated with a 0.2 standard deviation stronger perception that in the last year construction delays have occurred or projects have been abandoned in cases where advanced payment was received (listed in the figure as undesired project outcomes); this could reflect closer attention to this challenge as a result of GSAM. However, neither GSAM condition is associated with differences in the reported frequency with which district projects are abandoned, nor the amount of on-time project completions.

Outcome Family 2: Governance Quality: Turning to Outcome Family 2 and the measures of governance quality (Figure 11 below), neither CSO nor GAS programming are associated with significant changes in most outcomes, including perceptions of corruption in one's district compared to others; perceptions that corruption impedes development projects; or the degree of reported project transparency.

Figure 11: Governance Quality Effects among Administrators



Nevertheless, survey experimental evidence on the partisan manipulation of public resources does provide some suggestive evidence of an effect of GAS programming. There are reasons to believe that administrative and political officials have incentives to present an overly favorable picture of district development, i.e. that their responses to standard questions were subject to social desirability bias. Above and beyond standard questions, the endline surveys included a list survey experiment designed to provide information on the extent to which administrators see parties providing privileged access to public resources to their co-partisans. These survey experiments are particularly useful at uncovering the incidence of sensitive or unpopular behaviors or beliefs in a population—behaviors and beliefs that traditional survey questions are poor at uncovering. In the experiment, one half of the respondents were assigned to a control group, provided an innocuous list of factors that they might regularly see in their work, and asked to report the number of them they had actually seen. The other one half of respondents were provided the same list of options, albeit with the addition of an option for "Government officials providing members of their political party with privileged access to public resources." By comparing the means of the survey experimental treatment and control groups across GSAM treatment groups we are able to assess the incidence of partisan bias in the targeting of public resources. The results of the survey experiment fall just below statistical significance, but they suggest administrators in CSO districts are almost twice as likely to indicate that partisan bias is prevalent (39 percent vs. 17 percent in control districts) and more

than twice as likely (39 percent) in GAS districts. These mean differences are large, but they are not statistically significant given the large variance in responses. Nevertheless, these findings suggest that GSAM might have sensitized administrators to the possibility of partisan bias.

Outcome Family 3: Accountability: For accountability measures in Outcome 3 (Figure 12 and

Figure 13 below), GAS audits correspond with a stronger feeling that the lack of citizen participation in planning and oversight will represent a big obstacle to building better development projects. In CSO audit districts, administrators typically spend 3 more hours per week with citizens and responding to constituent concerns than their peers in control districts. These findings suggest an increased level of government responsiveness due to CSO interventions. GSAM has no effect on the perception that development project locations are chosen based on need.

Figure 12: Accountability Effects among Administrators

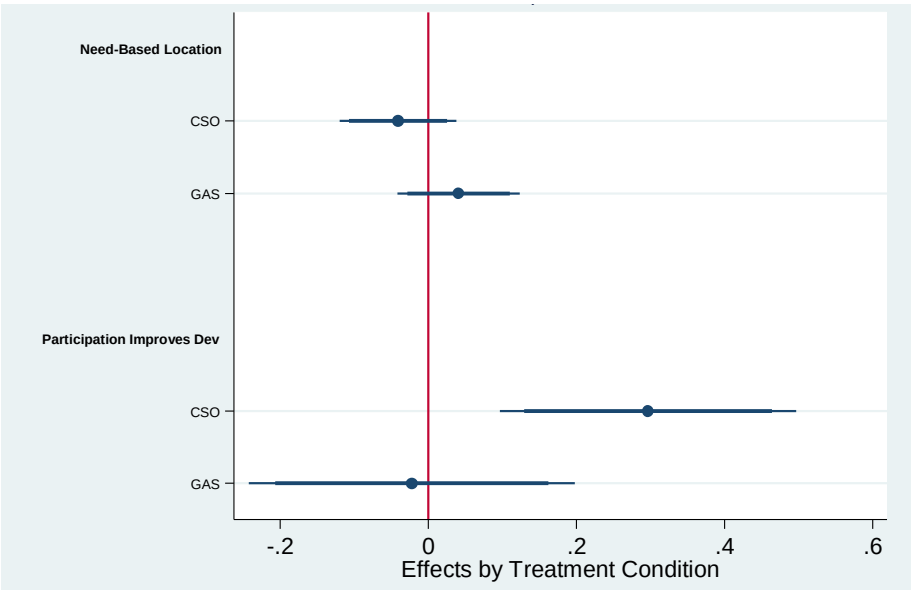
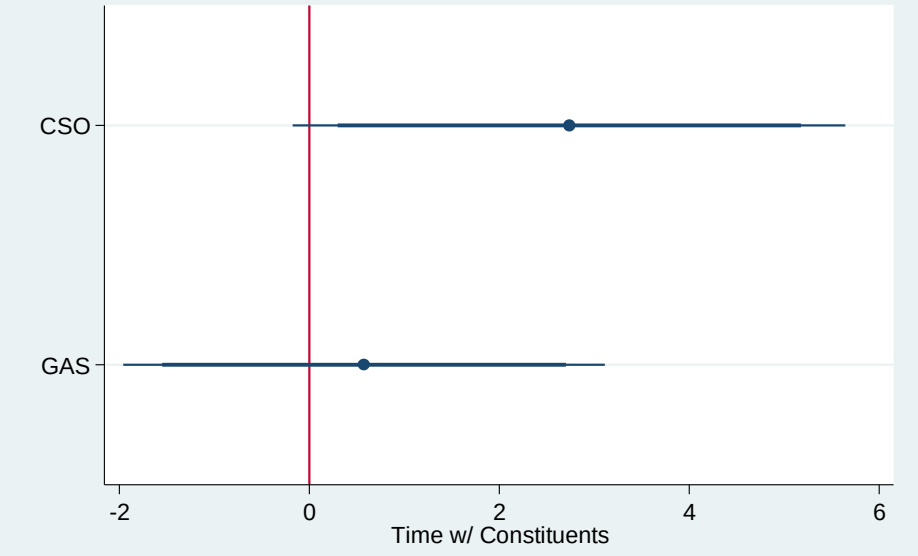


Figure 13: Time with Constituent Effects among Administrators

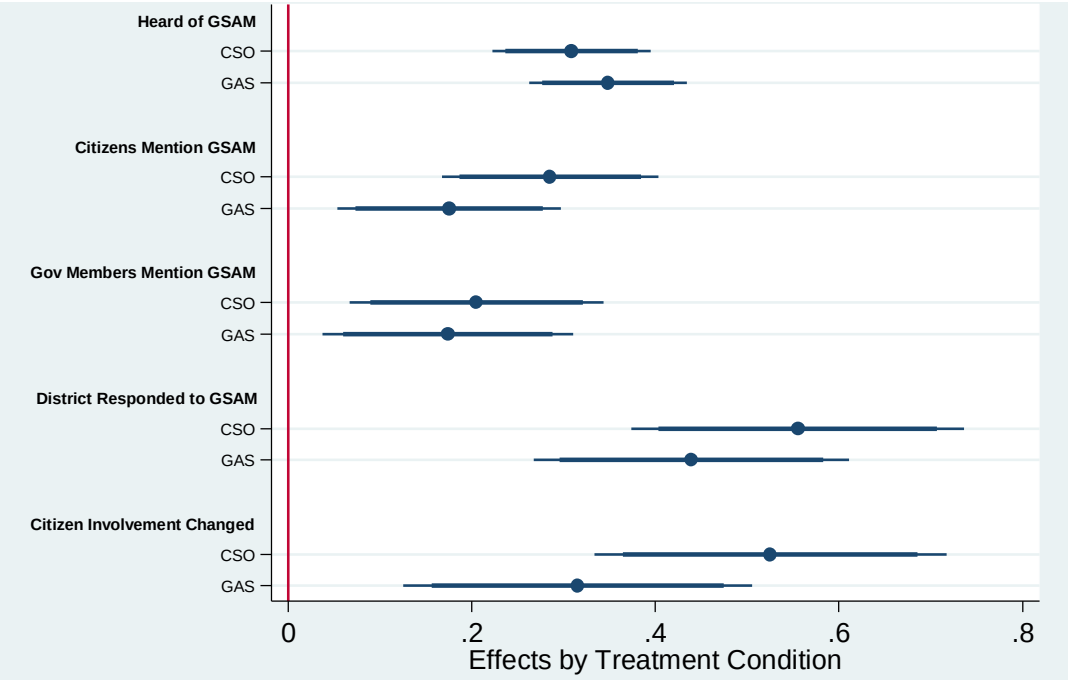


Analysis of Secondary Indicators: None of the secondary indicators for Outcomes 1 and 2 are significant for administrators, regardless of GSAM intervention type. See Appendix for full results.

7.4 POLITICIAN RESULTS

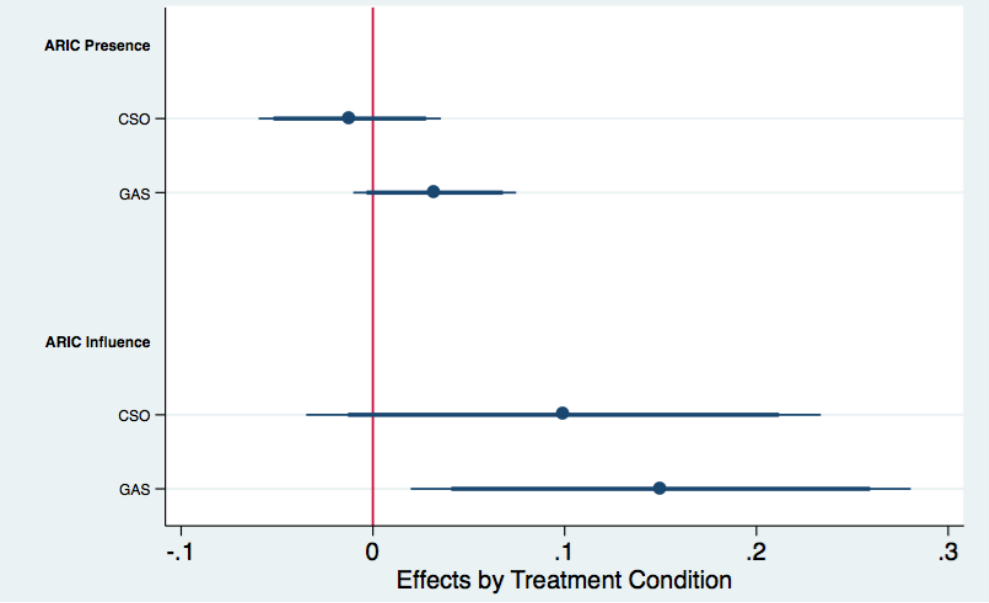
GSAM Awareness: Politicians in both CSO and GAS districts show a high rate of familiarity with GSAM (Figure 14 below). 68% of the politicians surveyed in CSO districts and 72% of politicians in GAS districts reported that they had heard of GSAM, compared to 37% of politicians in control districts. For the subset of politicians familiar with GSAM, we see significant, positive results on all of our indicators of GSAM engagement. GSAM programming corresponds with politicians being more likely to say that citizens have mentioned GSAM to them (20-30 percentage points higher); that other officials have mentioned GSAM to them (about 18-20 percentage points higher); that they have taken action in response to GSAM (45-55 percentage points higher); and that they perceive higher citizen engagement due to GSAM. For this last measure, citizen engagement due to GSAM was reported as being 30 percentage points higher in GAS districts and 50 percentage points higher in CSO districts.

Figure 14: GSAM Awareness and Engagement Effects among Politicians



Outcome Family 1: Development Project and Service Quality: Despite strong familiarity with GSAM in target districts, as well as perceived improvements in discussion of GSAM, district responsiveness, and citizen involvement due to the programming, few other outcomes (from the perspective of politician respondents) appear to be affected by GSAM programming. Politicians in GSAM districts are no more likely to report having an ARIC (Audit Report Implementation Committee) than control districts (Figure 15 below). However, among those who do report having an ARIC in the district, GSAM is associated with and increased perception of ARIC influence in GAS districts (at 90%).

Figure 15: Selected Development Project and Service Quality Effects among Politicians

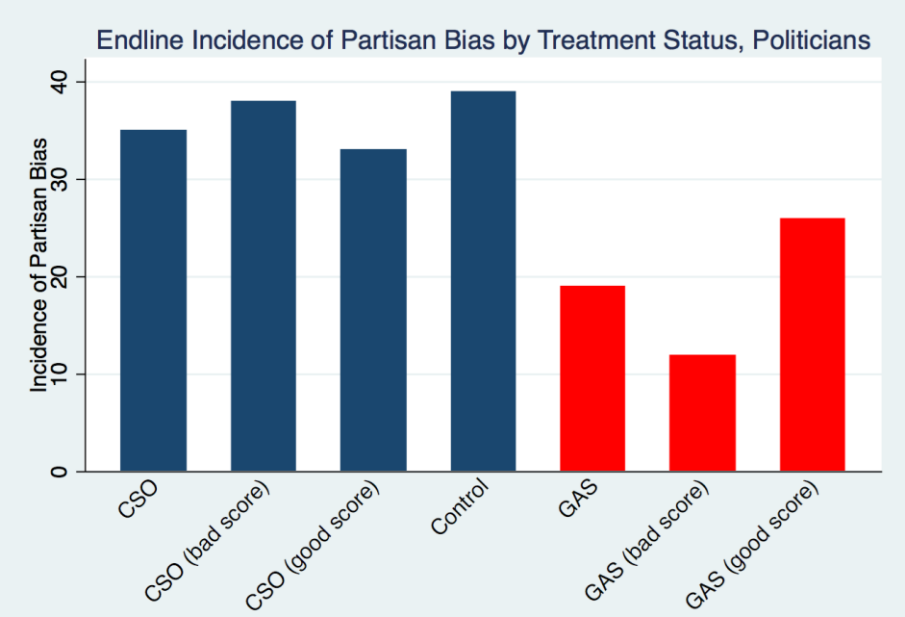


For other variables in Outcome 1 (capital projects and service delivery), GSAM does not show significant effects. Specifically, programming shows no significant association with reported project abandonment or project completion.

Outcome Family 2: Governance Quality: Similarly, we see no significant results for GSAM programming for Outcome 2, which includes: a lower perception of corruption in one's district compared to others; a greater belief corruption impedes development projects; and greater reported project transparency.

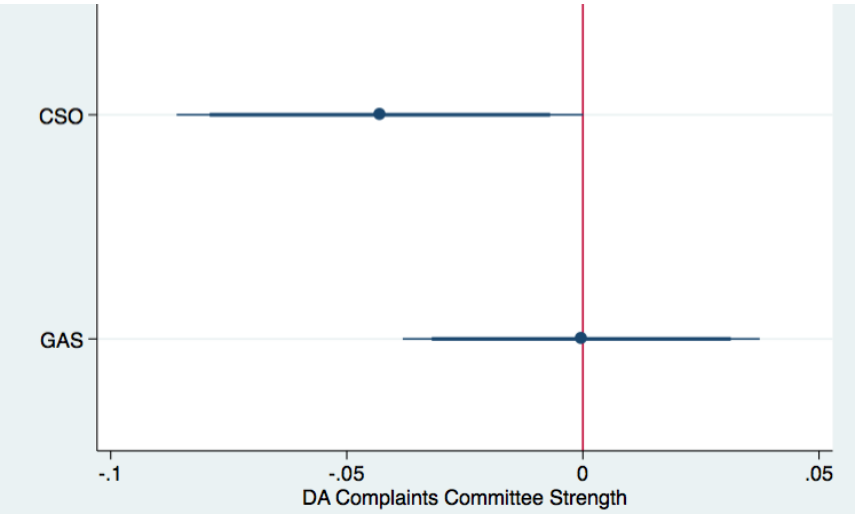
Turning to the survey experiment on partisan bias in access to public resources (see pages 38 and 39 above), we do find evidence that GAS programming sharply reduces the incidence of partisan targeting by politicians. Figure 16 provides a graphic presentation of the incidence of partisan bias by treatment arm and by whether or not the district received a good or bad audit score. The highlighted red results are statistically significant and show that GAS programming reduces the incidence of partisan bias in the use of public resources by half (from 39 percent of politicians to about 19 percent), and that the reduction is particularly stark in districts that received poor audit scores. This is an important substantive finding that also underscores the value of using survey experiments in contexts of sensitive issues.

Figure 16: Survey Experimental Evidence on Partisan Bias



Outcome Family 3: Accountability: For accountability measures in Outcome 3 (Figure 17), CSO audits correspond with a perception that DA Public Relations Complaints committees are less active. This could result from GSAM opening direct lines of communication between citizens, their DA members, and key members of the district administration; if so, the complaints committees could become redundant. GSAM has no effect on the perception that development project locations are chosen based on need, the amount of time spent with constituents, or the perceived role of citizen participation in building effective development projects.

Figure 17: Effects on Complaints Committee Strength Reported by Politicians



Analysis of Secondary Indicators: For politicians, we see no significant correlation between GSAM and any of our secondary outcomes (for capital projects and service delivery, governance quality, or accountability). When controls for years of experience and tribe are added to the analysis, though, we do find that GAS audits are

correlated with a greater perception that corruption plays an important role in officials losing elections (significant at 90%).

7.5 DISTRICT OUTCOME RESULTS

Figure 18 and Figure 19 present results on several objective measures of district management, including development project completion and the incidence of several forms of financial irregularities in district finances. As noted above, project non-completion is an important source of waste; if GSAM improves governance quality, it could be reflected in higher project completion rates. Likewise, if GSAM has improved DA management of budgets, of which development projects are a very large share, it could be reflected in fewer financial irregularities, particularly in the areas of contract irregularities and procurement irregularities that have direct bearing on development projects. All of these are measured for the most recent year available: the 2016 budget cycle. Financial irregularities are measured in Ghanaian cedis, which we normalize by the number of projects in the district.

Figure 18: District Outcomes: Project Completion

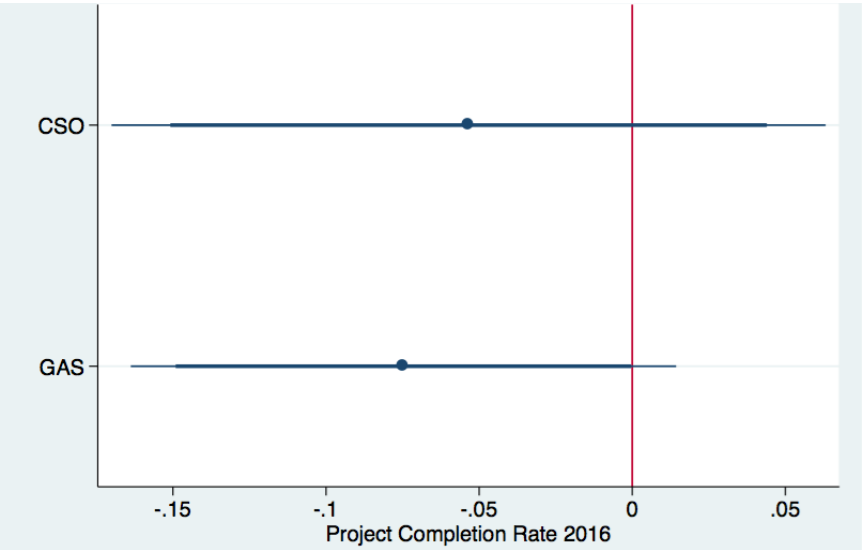


Figure 19: District Outcomes: Irregularities, Normalized by Project Count

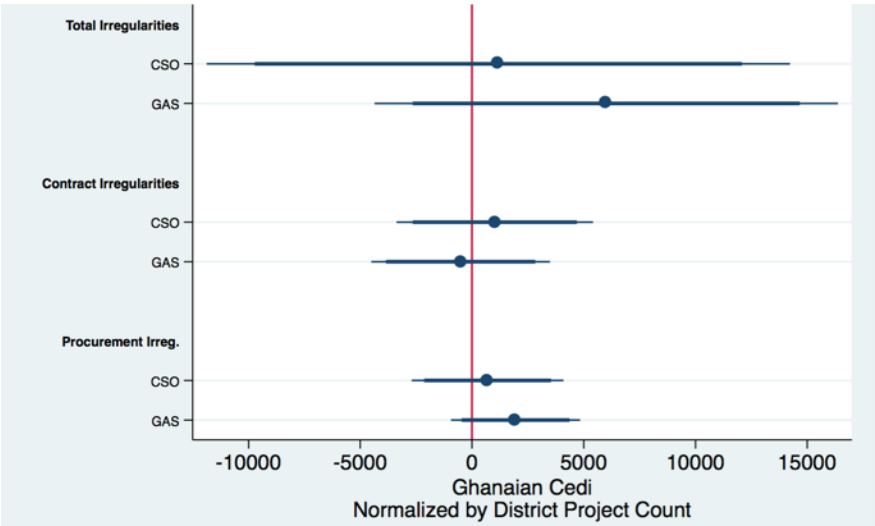


Figure 18 shows that neither CSO nor GAS programming has had an impact on project completion rates or financial irregularities. If anything, GAS programming may have slightly reduced project completion rates (at .10 significance level). Many district officials report considerable delays in receiving central government transfers (the District Assemblies Common Fund) and blame these delays on project non-completion. Since projects are contracted such that the DA makes payments to the contractor in phases over the life of projects, these delays in DACF funds can result in projects being suspended when contractors are not paid. Addressing this problem in Ghana's fiscal system is beyond GSAM, and it cannot be fixed by districts themselves.

8.0 CONCLUSIONS

8.1 SUMMARY OF FINDINGS

In this section we summarize the large body of findings and put them in a broader context. Table 1 provides a tabular summary of findings by hypothesis and indicator. Because the vast majority of non-null program effects identified through the household survey occur only in GSAM-intensive communities, we report these results in the table.

Table 1: Hypotheses and Conclusions

Outcome Measure	Data source	Conclusion
H1: Improved quality of capital projects and service delivery		
Increased access to development projects and corresponding services	Citizen survey	• GAS programming increases perceived benefits (to families and communities) from development projects. • CSO programming decreases perceived benefits from development projects. • Both CSO & GAS programming decrease satisfaction with the quality of construction.
Reduced obstacles to completing projects	Citizen survey	• No effects on perception of delays in either treatment group
Increased citizen satisfaction with services	Citizen survey	• CSO & GAS programming decrease satisfaction with services and project construction when audit results are poor.
Improved district planning for projects	Survey of officials, admin. data	• No effects on project completion rates.
Improved district contracting of projects	Survey of officials	• No effects on perceived abandonment in construction.
Improved district oversight of projects	Survey of officials, admin. data	• CSO programming increases the administrator-reported incidence of project delays, but has no effect on objective rates of project completion. It has no effect on politicians. • GAS programming increases the politician-reported influence of district Audit Report Implementation Committee; it has no effect on administrators.
H2: Improved quality of governance		
Improved perception of relative performance of district governments	Citizen survey	• No effects on perceptions of DA corruption compared with other DAs.

Outcome Measure	Data source	Conclusion
Reduced incidence of corruption	Citizen survey	• CSO programming has no effect on perceptions of corruption • GAS programming decreases citizen perception that DCEs govern in favor of their party or ethnic group at the expense of the public
Reduced incidence of partisan manipulation of public resources	Survey of officials	• Survey experimental evidence shows that GAS programming ***reduced*** partisan manipulation of public resources among politicians. Both GAS and CSO programming ***increased*** the perception of such manipulation by administrators.
Increased government transparency	Survey of officials	• No effects on officials' perceptions of transparency.
H3: Improved government accountability and citizen engagement		
Improved responsiveness to community needs in budgeting	Survey of officials	• CSO programming increases citizen-reported consultation on recent development projects; no effect on administrators or politicians reporting of participation opportunities. • GAS programming has no effect.
Improved responsiveness to community needs in project location	Survey of officials and citizen survey	• Among citizens, GAS programming may increase the perceived role of need in allocating projects, but not when audit results are negative. • CSO programming has no effect.
Increased accountability for how officials spend their time	Survey of officials	• In CSO districts, administrators spend three more hours per week responding to constituents; no effect on politicians. • GAS programming has no effect on officials' time allocations.
Increased accountability for how officials allocate projects	Survey of officials	• In both GAS and CSO districts, citizens are more likely to hold DA officials responsible for project quality. • In both GAS and CSO districts, citizens more likely to turn to an elected official for help with development projects.
Increased citizen knowledge and participation	Citizen survey and survey of officials	• No effects on citizens' reported participation. • GAS programming increases administrator reports that there is insufficient citizen engagement.
Knowledge & awareness of GSAM	Survey of officials	• No effects on citizens' awareness of GSAM • Large, positive effects on administrators and political officials' awareness of GSAM

In summary, CSO programming has generally had more effects on citizen than GAS programming, while GAS programming has had a bigger effect on governance quality. Across treatment arms, program effects have generally been stronger on citizen perceptions of district governments and their stated willingness to hold DA officials accountable, even if that has not been accompanied with more activism or participation. GSAM has

limited effects specifically on the governance of development projects, but GAS programming has encouraging effects on overall governance quality, particularly the incidence of partisan manipulation of public resources. The highlights bearing on each hypothesis are:

- **Quality of projects and services:** Both CSO and GAS programming generally reduce citizen satisfaction with projects and services, but this is largely driven by districts that receive negative audit reports. That citizens are correctly attributing bad audit performance to poor-performing DAs is encouraging from the point of view of accountability. This progress with citizens has not, however, translated into many substantial changes in how administrators or politicians manage projects or project budgets.
- **Quality of governance:** Neither GAS nor CSO programming improve transparency or corruption. GAS programming does reduce the incidence of partisan manipulation of public resources by politicians, and it also increases the perception of partisan manipulation among administrators. This is consistent with GAS sensitizing administrators to partisan manipulation and reducing its actual incidence among DA politicians.
- **Accountability:** GSAM has no effect on citizen participation in district governance, but both GAS and CSO programming increase the likelihood that citizens will hold DA officials accountable for project quality and are more likely to turn to a DA official for help with projects. CSO programming increases citizen-reported consultation on recent development, and administrators in CSO districts spend, on average, 3 hours more responding to constituents.

As discussed below, changing deep institutional practices bearing on project completion rates and financial irregularities will likely require more sustained institutional and fiscal change than is possible with a program like GSAM. Nevertheless, the overall results are consistent with some important improvements in accountability and project outcomes—citizens are sensitive to negative project outcomes and report increased willingness to hold DAs accountable. At the same time, DAs have somewhat reduced partisan manipulation of budgets and increased responsiveness to citizens. In short, GSAM has had some positive effects on the citizen-led, demand side of accountability, and some on the supply side of governance.

There is a growing body of research showing that top-down accountability initiatives, at least in the short term, can have a positive effect on the capacity of citizens to hold elected officials accountable. Evidence on bottom-up, citizen-led accountability is more mixed. Given the difficulty of linking citizen information, attitudinal change, engagement and participation to improved accountability, GSAM's effects on citizens are promising. This begs the question: why do we see stronger effects on citizens than some other, related impact evaluations? While other accountability-oriented governance projects focus on features of politics that are somewhat remote from citizens' day-to-day lives (such as, for instance, how elected officials spend their time), GSAM focuses on development projects that have very important, immediate implications for citizen wellbeing. Rural citizens care very much if they have a local school, a public toilet or a functional market. One key policy challenge is how to leverage GSAM's gains with citizens into broader institutional change in how DAs work.

This leads to a second question: why are the effects on administrators and political officials more modest? The weakness of the findings among DA officials, likely reflects several factors. First, GSAM has still reached a relatively small share of the broader DA citizenry. Thus, any pressure coming from the ground up is likely to be modest. Second, the underlying institutional capacity of DAs is limited by frequent administrator rotation, low reelection rates among political officials and scarce resources and skills for administering large development projects. These limitations are both political and administrative. Perhaps most importantly, they limit the incentives and capacity of DA officials to learn from, and respond to, citizens newly empowered by GSAM. Third, the system of fiscal transfers in Ghana makes DA budgeting and contracting difficult. Since central transfers are often delayed, contractors sometimes go unpaid, and projects are left incomplete. To the extent DA officials

cannot impact this binding fiscal constraint, this also limits their incentives to invest in being more responsive to citizens or institutional innovations that might improve DA institutional planning and oversight of projects.

8.2 POLICY RECOMMENDATIONS

This overarching interpretation of GSAM's effects suggest a number of policy recommendations.

1. GSAM should expand the extent of citizen outreach in CSO districts. Though programming is complete in GAS districts, it will continue for two more years in CSO districts. There are promising results in high-intensity CSO communities, but despite extensive efforts, GSAM is still only recognized by a small share of the citizenry, suggesting a need for additional outreach.
2. USAID and GSAM should develop and implement programming to strengthen the internal institutional capability of DAs. This should include development of an internal monitoring and evaluation system for development projects based on the weaknesses discovered in the CSO and GAS findings. Given the large sums involved in district development projects, there is a strong argument for digitizing planning, procurement and oversight of contracts.[22] In addition, programming could aim to strengthen training for key DA members and administrators involved in project planning and oversight, as well as professionalize the Audit Report Implementation Committees (ARICs) to increase their capacity to actively monitor DA budgeting and contracting of development projects. Given the turnover in these roles, this training would need to be ongoing and sustainable.
3. While the GSAM findings on GAS programming are modest, given that performance audits have been shown to be a useful tool for vertical oversight in other settings, we do not recommend discontinuing their use. Instead we recommend that USAID, the GAS, and the government of Ghana (GoG) explore and experiment with additional methods to increase audit impact based on the experiences of other countries. For example, some audit regimes require audited institutions to develop an action plan with an implementation calendar to address concerns raised by the audit. The central auditor then dispatches a follow-up team to verify if concerns have been addressed. Some audit regimes have further authorities to fine leaders for failing to develop or implement such plans. As above, audits might need to be paired with more extensive training of DA officials and/or more outreach to district citizens. As discussed below, these efforts can be combined with research aimed at understanding how best to maximize the influence of randomized audits in Ghana.
4. The GoG should improve the regularity of central transfers from Accra to the districts. Many DA officials cite the irregularity of central government transfers via the DACF (and other sources) and blame these delays on project non-completion. This results in delays of payment to contractors and suspension of project works. Project non-completion represents a serious waste of resources. Addressing this problem in Ghana's fiscal system is beyond GSAM, and it cannot be fixed by district governments or citizens themselves.
5. The GoG should reduce the rotation of district administrators. District administrators appear to rotate quite frequently. In our sample, 61% of administrators have been in their job two years or less, and 80% of them have been in their current district for four years or less. Short tenures can reduce the incentives of administrators to be responsive to local citizens and gather useful local knowledge. Short tenures can also reduce the incentives of administrators to address problems that require sustained investments of time and energy. Thus, the government of Ghana should consider reducing the frequency of rotation of administrators.
6. The GoG should consider efforts to professionalize DA and make careers as DA politicians more

[22] See the Construction Sector Transparency Initiative (CoST) for an example of one such initiative.

attractive. There is evidence from other settings that more professionalized legislatures—meaning those with more resources and longer tenures—produce better policy and oversight. Reelection rates to DAs are quite low (less than 50%), and 66% of political officials have two or fewer years of experience in the DA. While we do not fully understand why reelection rates are so low, they serve as a brake on accountability. Narrowly speaking, the turnover made it difficult for GSAM to impact politicians. More broadly, the weak reelection incentives probably reduce both the responsiveness of DA members to their constituents and their motivations to develop extensive knowledge about district budgeting, contracting and oversight. As such, the government of Ghana should consider steps to enhance the professionalism of DAs.

8.3 FUTURE POLICY-RELEVANT RESEARCH RECOMMENDATIONS

GSAM has pointed to a number of opportunities for future research on district governance. We briefly emphasize four of them here.

1. USAID and GSAM stakeholders should continue to assess the medium- to long-term impact of GSAM. Changing the quality of governance is often a slow process. GSAM's efforts to raise awareness and promote accountability might have larger effects over the long-term than over the short-term. Assessing the effects of the CSO programming two or three years from now is particularly important given that the programming will continue for several more years; indeed, there is a good argument for an additional "endline" for the CSO districts once the programming is done. It seems unlikely that the GAS programming will have lasting effects on citizens, but it very well could have a lasting impact on how administrators and political officials plan and implement development projects. In conjunction with points 2 and 3 below, it would be valuable to learn if and how exposure to GAS performance audit impacts administrators as they rotate positions and districts.
2. USAID and GSAM stakeholders should assess how best to leverage performance audit capacity. The GAS now has the internal capacity to conduct high-quality performance audits. Elsewhere such audits have been shown to be an effective tool for promoting accountability. Additional research, ideally in the context of field experiments, could be used to help address three key issues: (1) The ideal frequency and nature of performance audits that would promote better governance of district budgets at the lowest cost; (2) the most important complementary programming (e.g., training, follow-up policies, sanctioning/promotion) to promote better governance; and (3) the most cost-effective means of disseminating audit results to district officials and citizens to promote responsiveness.
3. Future research should seek to gain a better understanding of district-level political and administrative careers. To the extent GSAM or any other district-level programming or reforms are to have lasting impacts, they must consider the career trajectories and resulting incentives of district-level administrators and elected DA members. By most accounts, the Local Governance Service Secretariat's (LGSS) staff database is dated and incomplete. GSAM has collected some information on the background, training and recent careers of district officials. If combined with a carefully crafted online survey, this could provide a quick, inexpensive tool for key descriptive information. Ideally, this information would inform experimentation on: (1) alternative approaches to motivating hard work by district administrators (such experimentation is currently happening at the national level) and (2) different strategies for promoting more professional DAs.
4. Future research should also assess why some communities receive development projects while others do not. District budgets are very opaque, and it is unclear why DAs build projects in the places they do. Is it because some communities are more politically valuable, or because some have greater needs than others? The answer to this question is important since district-built projects can

have such important impacts on household wellbeing. To the extent that GSAM increases transparency and information, it might result in a planning process that is more likely to place projects in communities that have a greater need. In a complementary AidData-funded project, Wibbels, Ichino and Williams have geocoded approximately 45,000 projects from hard copies of annual progress reports (APRs) for 2011-2016 and combined them with very detailed census and electoral data, but this should be done on a permanent basis. If GSAM's effects are felt over the medium term, the impact on project location are likely to appear in 2017 and 2018 project allocation. The Principal Investigators have developed procedures for digitizing and geocoding APRs, so this can be done relatively efficiently and cheaply. Indeed, if made permanent, this could contribute to the development of a national infrastructure map, which the PIs are working with the GoG to produce.

REFERENCES

Andrabi, T., Das, J., and Khwaja, A. 2009. Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets. Unpublished manuscript.

Ayee, J. and Dickovick, JT. 2010. Comparative Assessment of Decentralization in Africa: Ghana Desk Study. Washington DC: United States Agency for International Development.

Björkman, M., and Svensson, J. 2009. Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. *The Quarterly Journal of Economics.* Vol. 124(2,1): 735-769.

CDD-Ghana 2008

Duflo, E., Hanna, R., and Ryan, S.P., 2012. Incentives Work: Getting Teachers to Come to School. *American Economic Review.* Vol. 102(4): 1241-78.

Ferraz, C. and Finan F. 2008. Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *The Quarterly Journal of Economics.* Vol. 123(2): 703-745.

Fox, L., Hoffman, B. Anyimadu, A., and Keshishian, M. 2011. Ghana Democracy and Governance Assessment: Final Report. Accra: United States Agency for International Development.

Ghana Audit Service. 2017. Report of the Auditor-General on the Management and Utilisation of District Assemblies' Common Fund and Other Statuatory Funds for the Year Ended 31 December 2016. Accra: Ghana Audit Service.

Honaker, J., and King, G., 2010. What to Do about Missing Values in Time-Series Cross-Section Data. *American Journal of Political Science.* Vol. 54(2): 561-581.

Huntington, H., Schultz, A., and Wibbels, E., 2015. Baseline Report for an Impact Evaluation of the G-SAM Project in Ghana. Washington DC: United States Agency for International Development.

Huntington, Heather, Erik Wibbels, and Brent Gall. 2017. Ghana Strengthening Accountability Mechanisms (GSAM) Impact Evaluation: Pre-Analysis Plan.

Kling, J., Liebman, J.B., and Katz, L.F., 2007. Experimental Analysis of Neighborhood Effects. *Econometrica.* Vol. 75(1): 83-119.

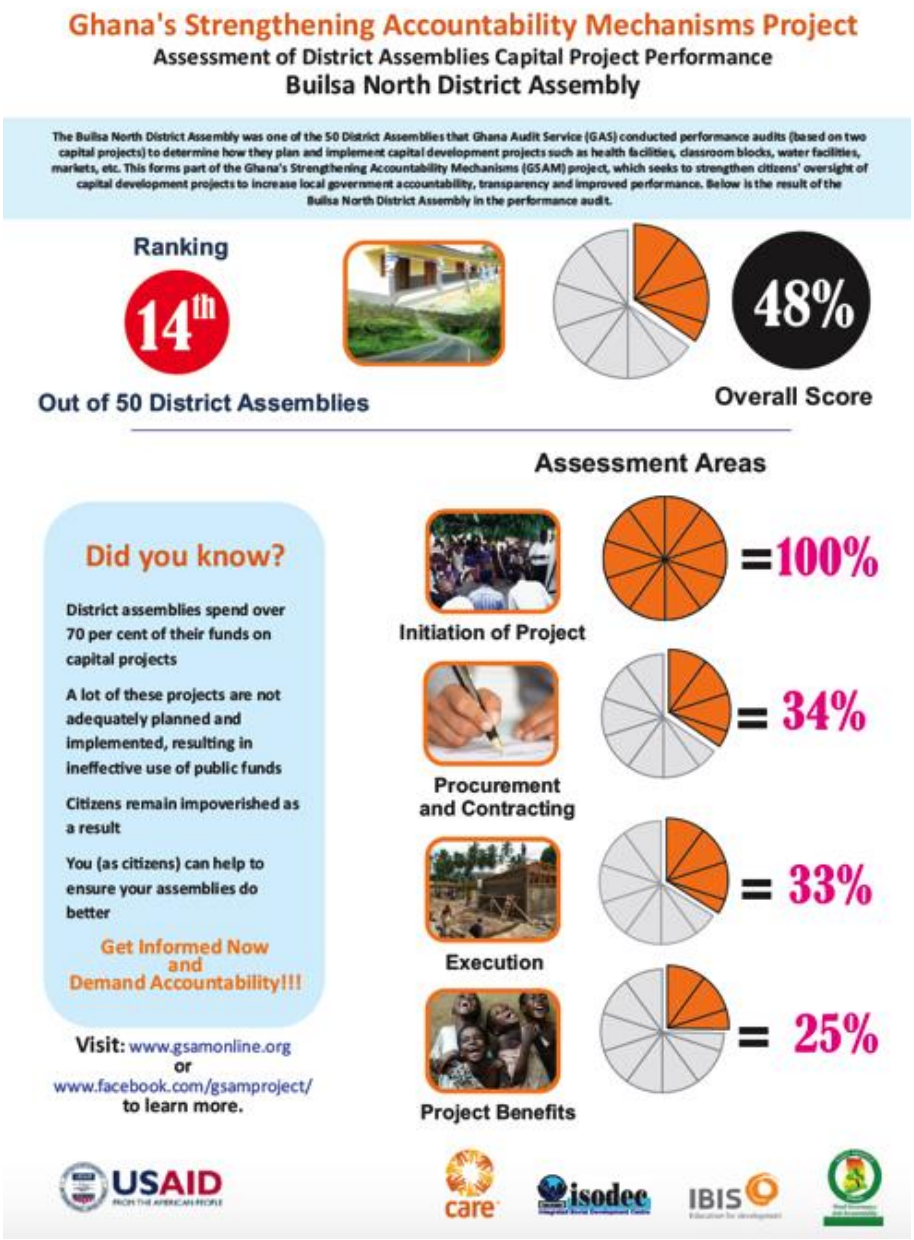
Muralidharan, K. and Sundararaman, V. 2011. Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy.* Vol. 119(1): 39-77.

Ohene-Konadu, G. 2001. *Gender Analysis and Interpretation of Barriers to Women's Participation in Ghana's Decentralised Local Government System.* Cape Town: African Gender Institute.

Olken, B. 2007. Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy.* Vol. 115(2): 200-249.

ANNEXES

ANNEX I: EXAMPLE OF SCORECARD FROM GAS PERFORMANCE AUDIT REPORT



ANNEX II: PRIMARY INDICATOR VARIABLES

Table 2: Indicators Derived from the Household Survey

Description	Hyp.	Survey Questions	Variable construction
CAPITAL PROJECTS AND SERVICE DELIVERY	**H1**		
Benefit from development projects		C4. How much has your FAMILY benefited from the project?	z-score of additive index: C4, C5
		C5. How much has your COMMUNITY benefited from the project?	
Satisfaction with school facilities		C27. Overall, how satisfied are you with the CONDITION OF THE FACILITIES at the BASIC school your child/children attend?	Ordinal: C27
		C28. How satisfied are you with the DISTANCE TO the BASIC school that your child/children attends?	
Perception of delays		READ: Samuel's district government began the construction of a school a year ago. Construction of the building was very slow, and no one has been seen working on it for six months. VIG2_3. How often do situations like those in Samuel's district happen in your district?	Ordinal: VIG2_3
Satisfaction with building		D7. How satisfied are you with how the District Assembly is BUILDING ROADS, or have you not heard enough to say?	z-score of additive index: D7, D8, D9
		D8. How satisfied are you with how the District Assembly is BUILDING BASIC HEALTH FACILITIES, or have you not heard enough to say?	
		D9. How satisfied are you with how the District Assembly is handling the BUILDING AND MAINTAINING OF SCHOOLS, or have you not heard enough to say?	
Satisfaction with services		D11. How satisfied are you with how the District Assembly is PROVIDING WATER, or have you not heard enough to say?	z-score of additive index: D11, D11a, D12, C9
		D11a. How satisfied are you with how the District Assembly is PROVIDING SANITATION SERVICES, or have you not heard enough to say?	
		D12. How satisfied are you with how the District Assembly is MAINTAINING LOCAL MARKET PLACES, or have you not heard enough to say?	
GOVERNANCE QUALITY	**H2**		
Governance obstacles to development		C56. How big of a problem is LACK OF CITIZEN PARTICIPATION AND OVERSIGHT to building better development projects?	z-score of additive index: C56, C58
		C58. How big of a problem is CORRUPTION AMONG DISTRICT PUBLIC OFFICIALS to building better development projects?	
Knowledge of district's relative performance		D14a. How common do you think GOVERNMENT CORRUPTION is in your district assembly compared to other district assemblies in Ghana?	Sum(Indicator if response to D14a =

			Don't Know, indicator if response to D14b = Don't Know)
		D14b. How common do you think INEFFICIENT DISTRICT GOVERNMENT SPENDING is in your district assembly compared to other district assemblies in Ghana?	
Perception of personal enrichment		D28. Sometimes officials benefit personally and financially from public projects. How much do you think your DCE benefits financially from the district's development projects, or haven't you heard enough about them to say?	z-score of additive index: D28, D29, D30
		D29. How much do you think YOUR ELECTED ASSEMBLY MEMBER benefits financially from the district's development projects? (or have you not heard enough to say?)	
		D30. How much do you think THE STAFF OF YOUR DA benefits financially from the district's development projects? (or have you not heard enough to say?)	
ACCOUNTABILITY	**H3**		
Project located based on community needs		C42. When you think of development projects in your district (such as schools, health clinics, electrification, and markets) how much do you think the NEEDS OF THE COMMUNITY influence where those development projects are located?	Ordinal: C42
Participation in project priority meetings		E2. Have you ever attended such a meeting?	Ordinal: E2
Opportunities for public input into development		C1. Please describe the largest development project that has been worked on in your community in the last two years…C3. Was your community consulted in the development of the project?	z-score of additive index: C3, E1, E3
		E1. Does your district assembly or town council ever hold public meetings to establish development priorities?	
		E3. How much influence do you think the people in this community have over decisions the District Assembly makes about development projects, such as school buildings, health clinics, irrigation ditches, or roads?	
GSAM AWARENESS			
Has heard of GSAM		G0. Have you heard of the Ghana Strengthening Accountability Mechanisms project? It is also referred to as G-SAM.	Binary: G0
Attended GSAM meeting		G2.Have you attended a community meeting organized by G-SAM?	Binary: G2
Contacted DA due to GSAM		G5. Have you contacted DA officials because of G-SAM?	Binary: G5
How did GSAM change involvement		G6. How has your involvement in your district's politics changed due to G-SAM?	Ordinal: G6
Overall GSAM evaluation		G7. How would you rate the G-SAM overall?	Ordinal: G7
Should GSAM continue		G8. Would you like G-SAM to continue as a program in your district?	Binary: G8

Table 3: Indicators Derived from the Administrator and Politician Survey

Description	Hyp.	Survey questions	Variable construction
CAPITAL PROJECTS AND SERVICE DELIVERY	**H1**		
Internal oversight		J1. Does your DA have an ARIC, that is, an Audit Report Implementation Committee?	z-score of additive index: J1, J2
		J2. Over the LAST YEAR how much influence would you say the ARIC has had on how the district contracts and oversees projects?	
Frequency of project abandonment		D20. In your estimation, what percentage of COMMON FUND PROJECTS are abandoned during construction (i.e. never finished) four years from when they began?	D20 and D21 collapsed to a single variable.
		D21. In your estimation, what percentage of DONOR FUNDED PROJECTS are abandoned during construction (i.e. never finished) four years from when they began?	
Project outcomes in past year		F1. How often have DELAYS IN CONSTRUCTION happened in THE LAST YEAR?	z-score of additive index: F1,F3,F11
		F3. How often in THE LAST YEAR have CONTRACTORS ABANDONED PROJECTS for which they have received advance payment?	
		F11. In your estimation, what percentage of large district capital projects have finished on time in THE LAST YEAR?	
Project completion rate		Share of projects completed in 2016; calculated from Annual Progress Reports	Percent
GOVERNANCE QUALITY	**H2**		
Relative district governance comparison		G1. How common do you think GOVERNMENT CORRUPTION is in your district assembly compared to other district assemblies in Ghana?	z-score of additive index: G1,G2
		G2. How common do you think INEFFICIENT DISTRICT GOVERNMENT SPENDING is in your district assembly compared to other district assemblies in Ghana?	
Corruption impeding development		G3. To what extent would you say that CORRUPTION is a problem in the PREPARATION of the Annual Action Plans?	z-score of additive index: G3,G4,H8,H9
		G4. To what extent would you say that CORRUPTION is a problem in the EXECUTION of the Annual Action Plans?	
		READ: There can be many obstacles to building development projects that serve the needs of district citizens. When you think about the preparation of the next Annual Action Plan to what extent do you agree that the following are likely to represent big obstacles to building better development projects in your district?...H8. To what extent do you agree or disagree that CORRUPTION AMONG DISTRICT PUBLIC OFFICIALS will represent a big obstacle?	
		READ: There can be many obstacles to building development projects that	

		serve the needs of district citizens. When you think about the preparation of the next Annual Action Plan to what extent do you agree that the following are likely to represent big obstacles to building better development projects in your district?....H9. To what extent do you agree or disagree that CORRUPTION AMONG CONTRACTORS will represent a big obstacle?	
Project transparency		H1. To what extent would you agree or disagree that district capital projects are PLANNED in a transparent manner?	z-score of additive index: H1,H2
		H2. To what extent would you agree or disagree that district capital projects are CONTRACTED in a transparent manner?	
ACCOUNTABILITY	**H3**		
Time spent on public service and receiving public input		[READ ALOUD] In a TYPICAL working week, how many hours do you spend on each of the following tasks? If you spend an hour working on something that belongs to more than one of these categories, please include it in both...B6d. Providing services/responding to citizen concerns	B6d + B6e
		[READ ALOUD] In a TYPICAL working week, how many hours do you spend on each of the following tasks? If you spend an hour working on something that belongs to more than one of these categories, please include it in both...B6e. Working with civil society or community groups	
Need as criteria for project location		D13. What is the SINGLE most important influence on where YOU want development projects to be located?	Recode D13 such that 5 (Targeting areas with the most need) = 1 and otherwise 0 (including Don't Know/Refused)
Perception that public participation will improve development?		H3. To what extent do you agree or disagree that CITIZENS' NON-PARTICIPATION IN PLANNING AND OVERSIGHT will represent a big obstacle to building better development projects?	Ordinal: H3
Strength of Public Complaints Committee		I3. Does the District Assembly have a Public Relations and Complaints Committee?	z-score of additive index: I3 and I4
		I4.How active is the DA Public Relations Complaints committee?	
GSAM AWARENESS			
Heard of GSAM		J14. Have you heard of the Ghana Strengthening Accountability Mechanisms project? It is also referred to as GSAM.	Binary: J14
Citizen mentions of GSAM		J16. Have citizens in your district mentioned the GSAM project to you OR others in your office?	Binary: J16
Government member mentions of GSAM		J19. Have other members of the district government mentioned the GSAM project to you or others in your office?	Binary: J19
Whether district responded to GSAM		J21. Has the district taken any action in response to GSAM?	Binary: J21

How district responded to GSAM		J22. What action has the district taken in response to the GSAM?	J22
Whether citizen involvement changed due to GSAM		J17. Has citizen involvement in your district's politics changed due to GSAM?	Ordinal: J17
How citizen involvement changed due to GSAM		J18. How has citizen involvement changed?	J18

ANNEX III: SECONDARY INDICATOR VARIABLES

The richness of our survey data allows for the creation of several alternative dependent variables for each of the outcome families described above. We describe these variables below; secondary analyses will be conducted on these outcome measures per Section 6 above.

OUTCOME FAMILY 1: CAPITAL PROJECTS AND SERVICE DELIVERY

Household data:

- Index of citizen access to services: how long it takes members of the household to walk to the nearest health facility, public toilet, and basic (elementary) school.
- Index of citizen satisfaction with the DA in providing: a) water, b) sanitation, and c) maintaining local market places.
- Continuous measure of the length of time it took to build the most recent school.
- Continuous measure of perceived importance of incomplete development projects.

Politician and Administrator Data:

- Ordinal measure of whether political considerations impact the choice of contractors (H5).
- A count of the number problems identified by respondents in the DAs systems for overseeing construction of development projects.
- Ordinal measure of the extent to which respondents agree that DA members gain votes by starting projects even if they aren't finished.
- Index of the extent to which respondents think members of the DA would rather complete a smaller number of projects rather than starting a larger number of them but leaving them incomplete.
- Index of deviations in contracting standards as measured by the frequency with which: a) fewer than the required number of contractors submitted bids and b) large projects were broken into smaller parts to avoid legal requirements for competitive bids.

OUTCOME FAMILY 2: DISTRICT GOVERNANCE QUALITY

Household data:

- Survey experimental evidence on the share of a 100,000 cedi DA windfall that would be wasted by the district government.
- Index of development-oriented political promises measured as the incidence of community development versus clientelistic benefits promised by politicians.
- Index of partisan favoritism in district governance as measured by citizen perceptions that: a) the DCE; and b) their DA representative will help members of their own party at the expense of the people of the district.
- Index of ethnic favoritism in district governance as measured by citizen perceptions that: a) the DCE; and b) their DA representative will help members of their own ethnic group at the expense of the people of the district.
- Ordinal measure of the importance of the relationship between local and national political parties to get development projects in respondents' community.
- Ordinal measure of the importance of being a member of the ruling party if someone in the respondent's village wants to get development projects built in their community.
- Ordinal measure of project fairness: Perception of inequity of project provision across communities in the district.

Politician and Administrator Data:

- The number of important changes that were made to the DA budget *after* it was passed.

OUTCOME FAMILY 3: CITIZEN PARTICIPATION AND DISTRICT ACCOUNTABILITY

Household Data:

- Index of the extent to which respondent holds DA, their elected Assembly member, and the DCE responsible for the quality of local school buildings.
- Indicator variable of whether respondent would turn to a member of the district government for help with their first development priority.
- Ordinal measure of satisfaction with the DCE.
- Ordinal measure of satisfaction with respondent's elected DA member.
- Index of whether respondent has filed a formal complaint with their DA member or the DA's Public Relations Complaint Committee.

Politician and Administrator Data:

- Index of the extent to which lack of responsiveness, unethical behavior, and corruption contribute to DA members losing election.

ANNEX IV: FULL HOUSEHOLD RESULTS

Annex 4.1: Main Results

Figure 20: Outcome 1 Program Effects

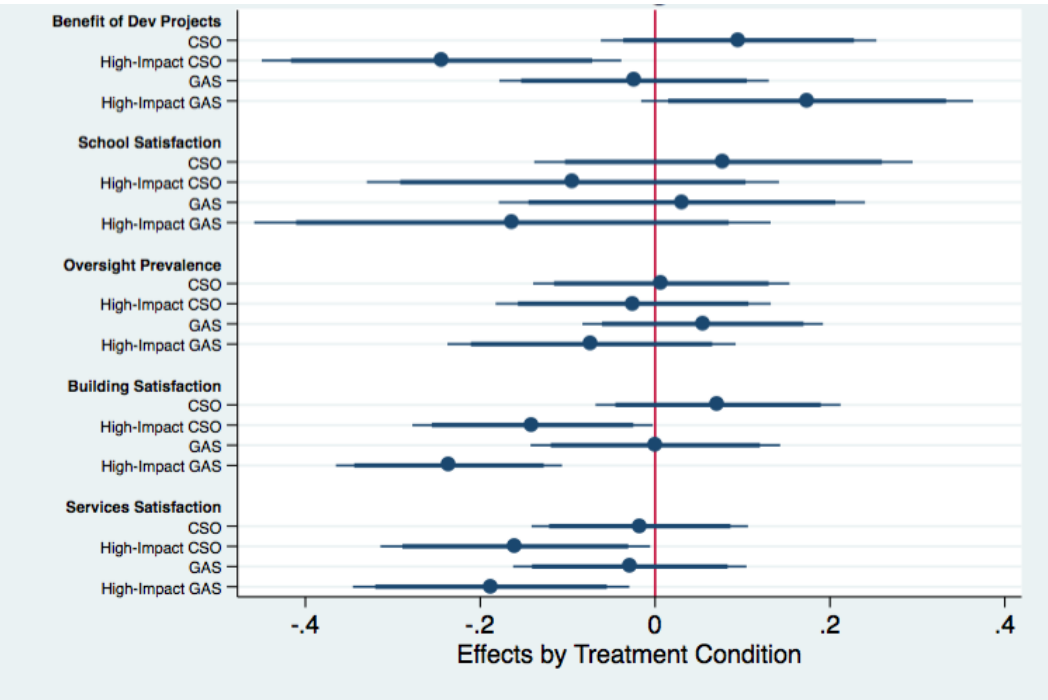


Table 4: Outcome Family 1

	(1) Benefit of Dev Projects	(2) School Satisfaction	(3) Oversight Prevalence	(4) Building Satisfaction	(5) Services Satisfaction
CSO	0.0955	0.0782	0.00707	0.0720	-0.0175
	(0.0798)	(0.110)	(0.0742)	(0.0710)	(0.0626)
High-Impact CSO	-0.244**	-0.0941	-0.0251	-0.140**	-0.160**
	(0.104)	(0.119)	(0.0796)	(0.0696)	(0.0780)
GAS	-0.0241	0.0307	0.0545	0.000282	-0.0289
	(0.0780)	(0.106)	(0.0696)	(0.0723)	(0.0676)
High-Impact GAS	0.174*	-0.163	-0.0727	-0.236***	-0.188**
	(0.0960)	(0.149)	(0.0834)	(0.0655)	(0.0800)
Constant	0.000468	2.658***	2.839***	-0.576**	-0.700***
	(0.191)	(0.318)	(0.166)	(0.221)	(0.166)
Observations	2746	2726	3888	3888	3888

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Figure 21: Outcome 2 Program Effects

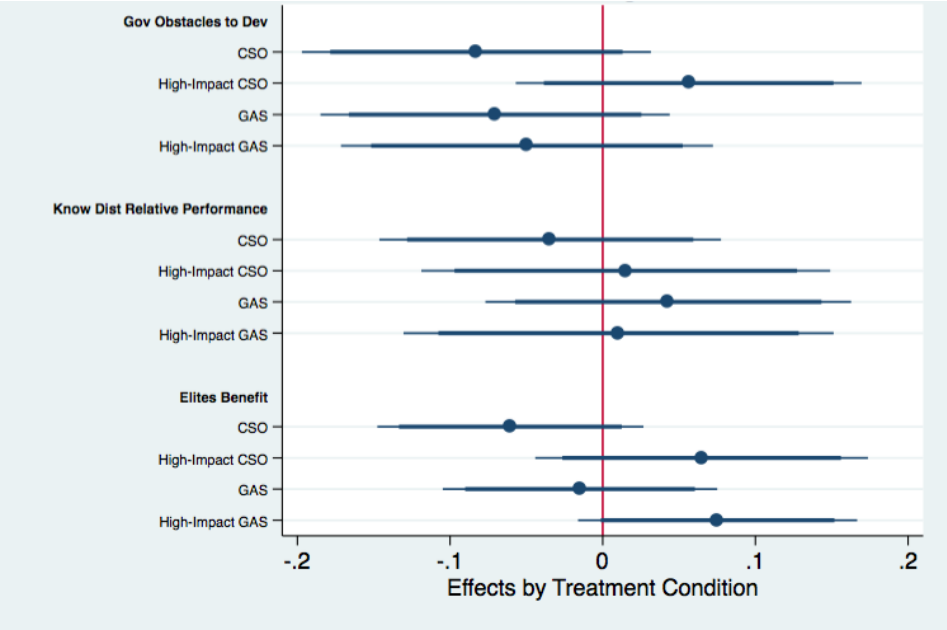


Table 5: Outcome Family 2

	(1) Gov Obstacles to Dev	(2) Know Dist Relative Performance	(3) Elites Benefit
CSO	-0.0827	-0.0344	-0.0605
	(0.0578)	(0.0566)	(0.0440)
High-Impact CSO	0.0563	0.0151	0.0648
	(0.0573)	(0.0678)	(0.0550)
GAS	-0.0705	0.0429	-0.0149
	(0.0579)	(0.0606)	(0.0454)
High-Impact GAS	-0.0496	0.0104	0.0753
	(0.0616)	(0.0713)	(0.0461)
Constant	0.149	0.956***	0.669***
	(0.190)	(0.0403)	(0.207)
Observations	3888	3888	3888

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Figure 22: Outcome 3 Program Effects

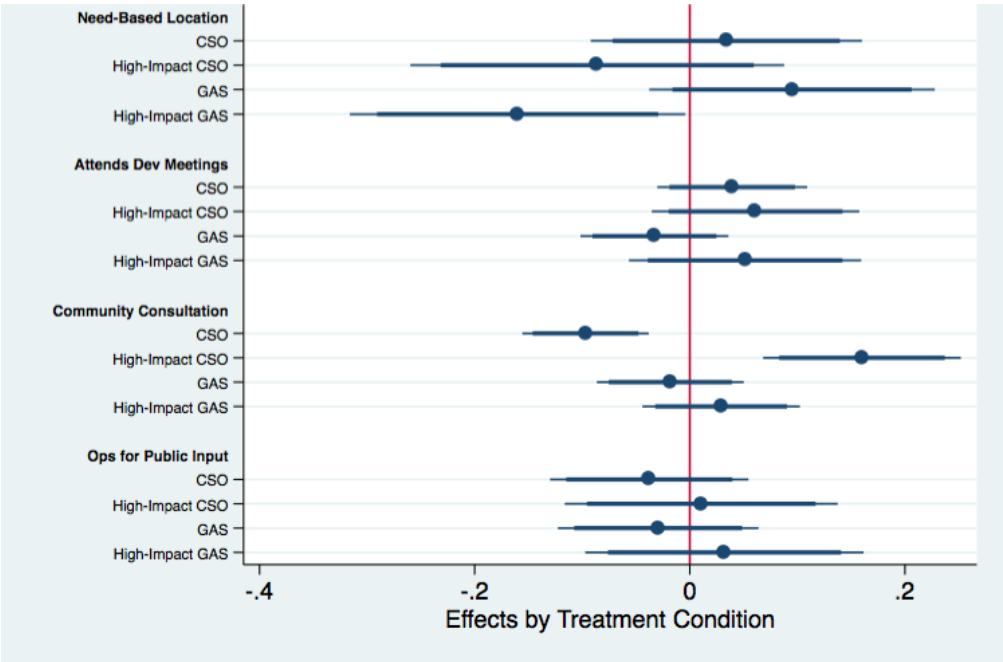


Table 6: Outcome Family 3

	(1) Need-Based Location	(2) Attended Project Mtg	(3) Community Consultation	(4) Ops for Public Input
CSO	0.0339	0.0394	-0.0969***	-0.0377
	(0.0638)	(0.0353)	(0.0297)	(0.0466)
High-Impact CSO	-0.0861	0.0612	0.160***	0.0106
	(0.0879)	(0.0488)	(0.0465)	(0.0642)
GAS	0.0951	-0.0329	-0.0181	-0.0294
	(0.0672)	(0.0348)	(0.0345)	(0.0472)
High-Impact GAS	-0.160**	0.0514	0.0291	0.0322
	(0.0789)	(0.0546)	(0.0370)	(0.0654)
Constant	3.045***	0.571***	0.783***	0.605**
	(0.0548)	(0.0814)	(0.0700)	(0.285)
Observations	3888	1872	2391	3888

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Figure 23: Program Effects on GSAM Engagement

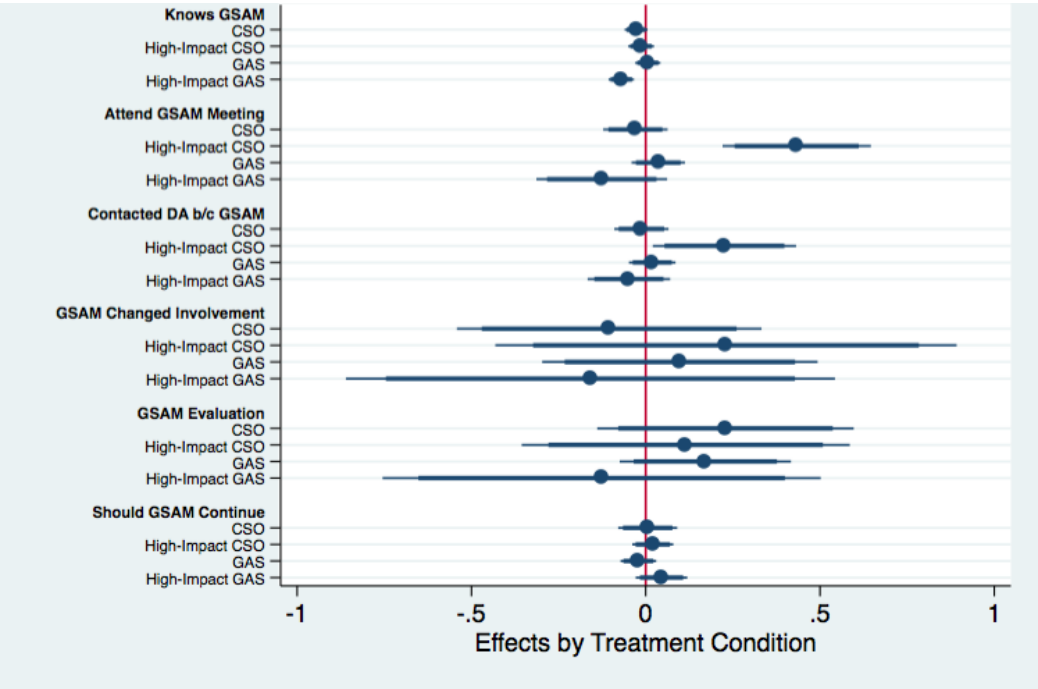


Table 7: GSAM Engagement

	(1) Knows GSAM	(2) Attend GSAM Meeting	(3) Contacted DA b/c GSAM	(4) GSAM Changed Involvement	(5) Services Satisfaction	(6) Should GSAM Continue
CSO	-0.0221	-0.00755	-0.0114	-0.270	0.274	0.0108
	(0.0151)	(0.0527)	(0.0438)	(0.225)	(0.206)	(0.0430)
High-Impact CSO	-0.00805	0.428***	0.197*	0.120	0.00833	0.0391
	(0.0177)	(0.115)	(0.101)	(0.305)	(0.268)	(0.0391)
GAS	0.0103	0.0496	0.0312	-0.0820	0.162	-0.0308
	(0.0173)	(0.0448)	(0.0387)	(0.209)	(0.156)	(0.0284)
High-Impact GAS	-0.0694***	-0.153*	-0.0643	-0.0486	-0.167	0.0629
	(0.0186)	(0.0796)	(0.0716)	(0.369)	(0.332)	(0.0435)
Constant	0.136**	-0.183	-0.112	3.792***	1.739***	0.925***
	(0.0580)	(0.296)	(0.120)	(0.741)	(0.523)	(0.118)
Observations	3815	322	316	249	240	278

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Annex 4.2: Household Outcomes with Controls

Table 8: Outcome Family 1

	(1) Benefit of Dev Projects	(2) School Satisfaction	(3) Oversight Prevalence	(4) Building Satisfaction	(5) Services Satisfaction
CSO	0.104	0.110	0.0143	0.0595	-0.0263
	(0.0804)	(0.104)	(0.0740)	(0.0647)	(0.0623)
High-Impact CSO	-0.262**	-0.0671	-0.0224	-0.108	-0.137*
	(0.105)	(0.119)	(0.0770)	(0.0673)	(0.0772)
GAS	-0.0209	0.0184	0.0246	-0.0521	-0.0336
	(0.0794)	(0.102)	(0.0711)	(0.0685)	(0.0699)
High-Impact GAS	0.177*	-0.155	-0.0631	-0.201***	-0.186**
	(0.0966)	(0.149)	(0.0824)	(0.0670)	(0.0802)
Constant	0.00586	2.173***	2.633***	-1.112***	-0.858***
	(0.313)	(0.376)	(0.246)	(0.239)	(0.252)
Observations	2746	2726	3888	3888	3888

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Table 9: Outcome Family 2

	(1) Gov Obstacles to Dev	(2) Know Dist Relative Performance	(3) Elites Benefit
CSO	-0.0758	-0.0253	-0.0421
	(0.0568)	(0.0565)	(0.0410)
High-Impact CSO	0.0546	0.0190	0.0743
	(0.0570)	(0.0665)	(0.0513)
GAS	-0.0856	0.0701	-0.0247
	(0.0579)	(0.0609)	(0.0418)
High-Impact GAS	-0.0293	-0.0262	0.0862*
	(0.0616)	(0.0736)	(0.0448)
Constant	0.428**	0.724***	0.954***
Observations	3888	3888	3888

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Table 10: Outcome Family 3

	(1) Need-Based Location	(2) Attended Project Mtg	(3) Community Consultation	(4) Ops for Public Input
CSO	0.0344	0.0302	-0.103***	-0.0377
	(0.0641)	(0.0316)	(0.0290)	(0.0462)
High-Impact CSO	-0.100	0.0578	0.147***	0.0195
	(0.0873)	(0.0455)	(0.0467)	(0.0645)
GAS	0.0750	-0.0299	-0.0177	-0.0258
	(0.0685)	(0.0327)	(0.0327)	(0.0472)
High-Impact GAS	-0.167**	0.0616	0.0179	0.0405
	(0.0797)	(0.0509)	(0.0372)	(0.0635)
Constant	2.727***	0.818***	0.871***	0.749***
	(0.199)	(0.115)	(0.134)	(0.246)
Observations	3888	1872	2391	3888

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 11: GSAM Engagement

	(1) Knows GSAM	(2) Attend GSAM Meeting	(3) Contacted DA b/c GSAM	(4) GSAM Changed Involvement	(5) Services Satisfaction	(6) Should GSAM Continue
CSO	-0.0221	-0.00755	-0.0114	-0.270	0.274	0.0108
	(0.0151)	(0.0527)	(0.0438)	(0.225)	(0.206)	(0.0430)
High-Impact CSO	-0.00805	0.428***	0.197*	0.120	0.00833	0.0391
	(0.0177)	(0.115)	(0.101)	(0.305)	(0.268)	(0.0391)
GAS	0.0103	0.0496	0.0312	-0.0820	0.162	-0.0308
	(0.0173)	(0.0448)	(0.0387)	(0.209)	(0.156)	(0.0284)
High-Impact GAS	-0.0694***	-0.153*	-0.0643	-0.0486	-0.167	0.0629
	(0.0186)	(0.0796)	(0.0716)	(0.369)	(0.332)	(0.0435)
Constant	0.136**	-0.183	-0.112	3.792***	1.739***	0.925***
	(0.0580)	(0.296)	(0.120)	(0.741)	(0.523)	(0.118)
Observations	3815	322	316	249	240	278

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Annex 4.3: Household Outcomes: High vs. Low Audit Scores

Table 12: Outcome Family 1 High Audit Rank

	(1) Benefit of Dev Projects	(2) School Satisfaction	(3) Oversight Prevalence	(4) Building Satisfaction	(5) Services Satisfaction
CSO	0.269**	0.0611	0.212*	0.172	0.0673
	(0.119)	(0.170)	(0.111)	(0.120)	(0.0891)
High-Impact CSO	-0.242*	-0.0831	-0.133	-0.155	-0.0926
	(0.142)	(0.189)	(0.120)	(0.117)	(0.115)
High-Impact GAS	0.259**	-0.185	-0.0172	-0.251**	-0.127
	(0.123)	(0.231)	(0.117)	(0.104)	(0.120)
Constant	0.00735	2.749***	3.106***	-0.485***	-0.691***
	(0.0324)	(0.0839)	(0.0390)	(0.0359)	(0.0421)
Observations	1034	1084	1492	1492	1492

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 13: Outcome Family 1 Low Audit Rank

	(1) Benefit of Dev Projects	(2) School Satisfaction	(3) Oversight Prevalence	(4) Building Satisfaction	(5) Services Satisfaction
CSO	0	0.456***	0.104	0.325***	0.296**
	(.)	(0.154)	(0.0977)	(0.0864)	(0.123)
High-Impact CSO	-0.361**	-0.227	-0.00219	-0.197***	-0.300***
	(0.155)	(0.148)	(0.107)	(0.0698)	(0.104)
High-Impact GAS	0.176	-0.145	-0.156	-0.267***	-0.240**
	(0.173)	(0.161)	(0.135)	(0.0681)	(0.116)
Constant	0.533***	1.744***	2.406***	-1.243***	-1.287***
	(0.0729)	(0.144)	(0.121)	(0.0872)	(0.118)
Observations	971	933	1354	1354	1354

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 14: Outcome Family 2 High Audit Rank

	(1)	(2)	(3)
	Gov Obstacles to Dev	Know Dist Relative Performance	Elites Benefit
CSO	0.146	0.0132	0.116*
	(0.0966)	(0.0794)	(0.0677)
High-Impact CSO	0.0600	0.00645	-0.0419
	(0.0836)	(0.0929)	(0.0780)
High-Impact GAS	-0.0263	-0.0663	0.0578
	(0.0886)	(0.0992)	(0.0554)
Constant	0.0308	1.022***	0.429***
	(0.0367)	(0.0331)	(0.0374)
Observations	1492	1492	1492

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 15: Outcome Family 2 Low Audit Rank

	(1)	(2)	(3)
	Gov Obstacles to Dev	Know Dist Relative Performance	Elites Benefit
CSO	-0.0567	-0.0703	-0.0677
	(0.0891)	(0.139)	(0.0686)
High-Impact CSO	-0.0164	-0.0148	0.101
	(0.0780)	(0.105)	(0.0724)
High-Impact GAS	-0.0728	0.114	0.0698
	(0.0897)	(0.113)	(0.0719)
Constant	0.506***	1.042***	1.105***
	(0.0875)	(0.139)	(0.0667)
Observations	1354	1354	1354

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 16: Outcome Family 3 High Audit Rank

	(1) Need-Based Location	(2) Attended Project Mtg	(3) Community Consultation	(4) Ops for Public Input
CSO	-0.0842	0	-0.167**	0.0490
	(0.106)	(.)	(0.0632)	(0.0677)
High-Impact CSO	0.0530	0.0229	0.194***	-0.101
	(0.124)	(0.0869)	(0.0676)	(0.100)
High-Impact GAS	-0.0969	0.0210	0.0760	-0.0900
	(0.130)	(0.0823)	(0.0543)	(0.0900)
Constant	3.199***	0.438***	0.868***	0.187***
	(0.0435)	(0.0562)	(0.0151)	(0.0360)
Observations	1492	694	930	1492

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 17: Outcome Family 3 Low Audit Rank

	(1) Need-Based Location	(2) Attended Project Mtg	(3) Community Consultation	(4) Ops for Public Input
CSO	0.0864	0.121**	-0.00237	-0.215**
	(0.143)	(0.0552)	(0.0487)	(0.0844)
High-Impact CSO	-0.206	0.0642	0.152**	0.0818
	(0.127)	(0.0462)	(0.0701)	(0.0872)
High-Impact GAS	-0.262***	0.123	-0.0213	0.272***
	(0.0952)	(0.0826)	(0.0593)	(0.0874)
Constant	2.982***	0.621***	0.521***	1.357***
	(0.141)	(0.0509)	(0.0469)	(0.0944)
Observations	1354	681	829	1354

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Annex 4.4: Household Outcomes: Secondary Indicators

Table 18: Outcome Family 1 Secondary Indicators

	(1)	(2)
	Walking Distance to Services	Dev Project Importance
CSO	-0.00266	-0.0603**
	(0.0646)	(0.0301)
High-Impact CSO	0.0819	0.0112
	(0.103)	(0.0291)
GAS	-0.0512	-0.0313
	(0.0576)	(0.0287)
High-Impact GAS	0.210**	-0.0166
	(0.0941)	(0.0337)
Constant	0.227	1.014***
	(0.155)	(0.0267)
Observations	3888	3888

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 19: Outcome Family 2 Secondary Indicators

	(1)	(2)	(3)	(4)	(5)	(6)
	Expected Windfall Waste	Community vs Clientelism	Party Favoritism	Ethnic Favoritism	Gov Connections Needed	Ruling Party Needed
CSO	-1.663	0.154	-0.0285	-0.0392	-0.0101	-0.0221
	(1.466)	(0.118)	(0.0533)	(0.0587)	(0.0639)	(0.0733)
High-Impact CSO	4.541**	0.219*	-0.0310	-0.0257	-0.111	0.00168
	(1.932)	(0.122)	(0.0738)	(0.0611)	(0.0742)	(0.0773)
GAS	-1.535	0.0678	-0.110*	-0.142**	-0.00547	-0.0647
	(1.592)	(0.121)	(0.0565)	(0.0639)	(0.0657)	(0.0755)
High-Impact GAS	1.390	-0.0777	0.00959	0.0987	-0.119	-0.00126
	(2.128)	(0.144)	(0.0590)	(0.0691)	(0.0762)	(0.0819)
Constant	37.17***	0.0750	0.361***	0.712***	4.725***	4.707***
	(5.297)	(0.716)	(0.0801)	(0.0814)	(0.140)	(0.290)
Observations	3888	3888	3888	3888	3888	3888

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 20: Outcome Family 3 Secondary Indicators

	(1) Pols Held Responsible	(2) Turn to Gov for Help	(3) DCE Satisfaction	(4) DA Satisfaction	(5) Filed Complaint
CSO	0.117*	0.0641**	-0.0259	0.0273	-0.00150
	(0.0612)	(0.0272)	(0.0951)	(0.0928)	(0.0169)
High-Impact CSO	0.0699	-0.0470	0.0274	-0.173	0.0107
	(0.0661)	(0.0286)	(0.135)	(0.134)	(0.0253)
GAS	0.139**	0.0529*	0.134	0.161	-0.00566
	(0.0615)	(0.0286)	(0.104)	(0.105)	(0.0159)
High-Impact GAS	-0.125*	-0.0103	-0.122	-0.180	0.0249
	(0.0702)	(0.0369)	(0.124)	(0.126)	(0.0274)
Constant	0.130	0.676***	4.664***	5.829***	0.0404*
	(0.173)	(0.0848)	(0.0722)	(0.129)	(0.0217)
Observations	3888	3888	3888	3888	3888

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

ANNEX V: ADMINISTRATOR RESULTS

Annex 5.1: Main Results

Figure 24: Outcome 1, Administrators

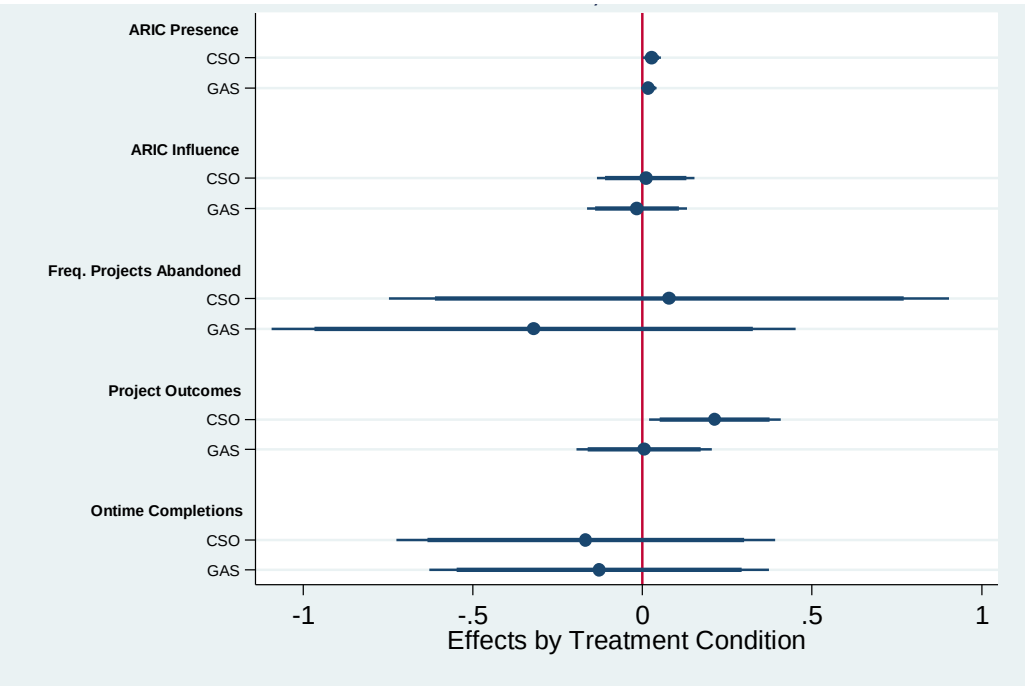


Table 7: Outcome Family 1

	(1) ARIC Presence	(2) ARIC Influence	(3) Freq. Projects Abandoned	(4) Project Outcomes	(5) Ontime Completions
CSO	0.0280**	0.0100	0.0789	0.213**	-0.167
	(0.0130)	(0.0724)	(0.417)	(0.0982)	(0.282)
GAS	0.0188	-0.0160	-0.320	0.00534	-0.128
	(0.0116)	(0.0744)	(0.390)	(0.101)	(0.254)
Constant	0.984***	2.469***	1.347**	-0.0944	8.165***
	(0.0104)	(0.161)	(0.586)	(0.166)	(0.573)
Observations	739	670	760	760	666

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Figure 25: Outcome 2, Administrators

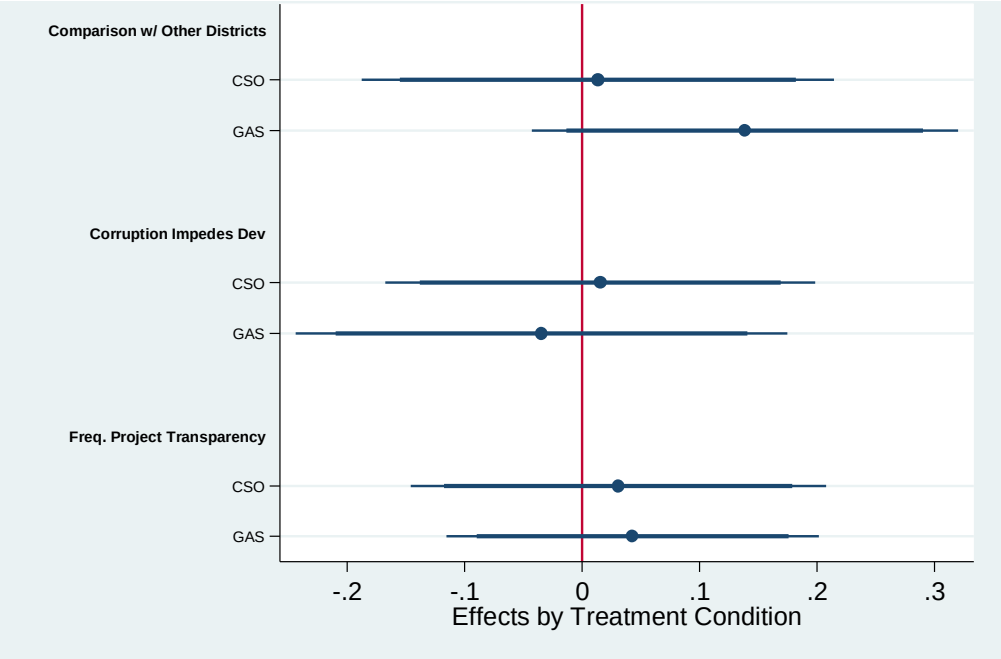


Table 21: Outcome Family 2

	(1) Comparison w/ Other Districts	(2) Corruption Impedes Dev	(3) Project Transparency
CSO	0.0133	0.0153	0.0307
	(0.102)	(0.0926)	(0.0895)
GAS	0.138	-0.0348	0.0428
	(0.0917)	(0.106)	(0.0802)
Constant	-0.319***	0.369**	0.244**
	(0.0776)	(0.148)	(0.118)
Observations	760	760	760

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Figure 26: Outcome 3, Administrators

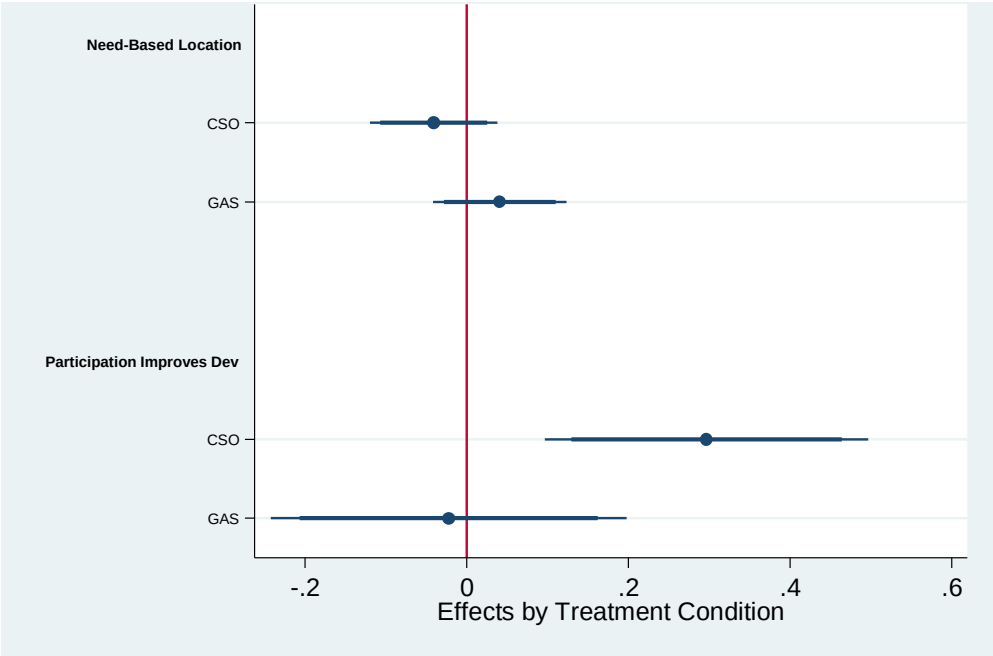


Figure 27: Outcome 3, Time with Constituents Administrators

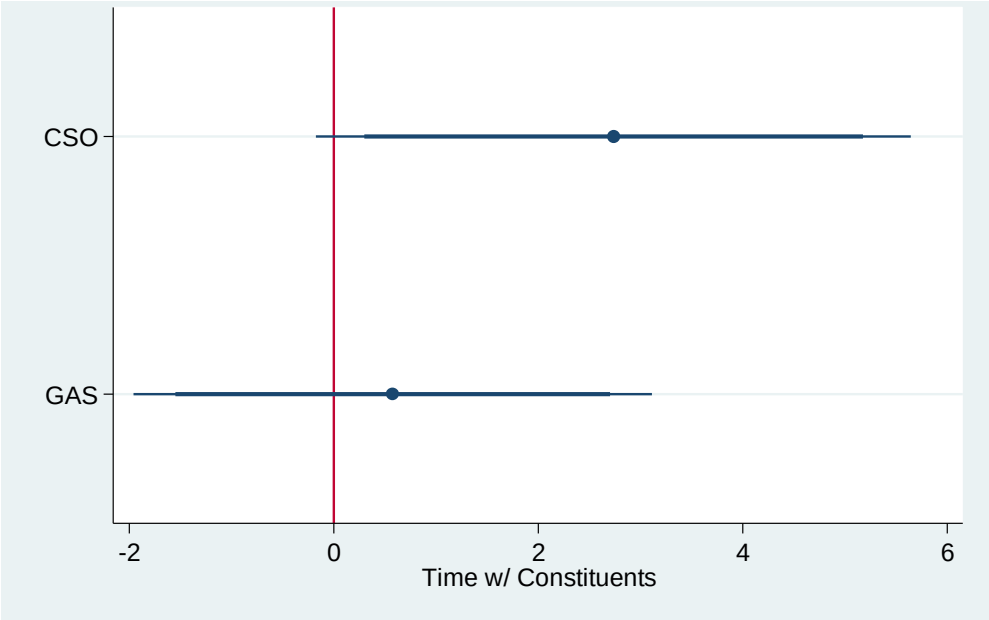


Table 22: Outcome Family 3

	(1) Participation Improves Dev	(2) Time w/ Constituents	(3) Need-Based Location
CSO	0.297***	2.737*	-0.0407
	(0.101)	(1.473)	(0.0400)
GAS	-0.0222	0.578	0.0410
	(0.111)	(1.284)	(0.0417)
Constant	4.442***	34.30***	0.800***
	(0.276)	(7.877)	(0.154)
Observations	760	760	732

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Table 23: GSAM Indicators

	(1) Heard of GSAM	(2) Citizens Mention GSAM	(3) Gov Members Mention GSAM	(4) District Responded to GSAM	(5) Citizen Involvement Changed
CSO	0.360***	0.356***	0.355***	0.449***	0.590***
	(0.0352)	(0.0587)	(0.0576)	(0.0677)	(0.154)
GAS	0.295***	0.201***	0.260***	0.289***	0.409**
	(0.0365)	(0.0603)	(0.0598)	(0.0662)	(0.159)
Constant	0.582***	0.118	0.182*	0.550***	4.160***
	(0.0393)	(0.0845)	(0.0927)	(0.0680)	(0.129)
Observations	755	480	489	353	314

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Annex 5.2: Administrator Outcomes with Controls

Table 24: Outcome Family 1

	(1) ARIC Presence	(2) ARIC Influence	(3) Freq. Projects Abandoned	(4) Project Outcomes	(5) Ontime Completions
CSO	0.0224**	0.00498	0.0780	0.223**	-0.196
	(0.0101)	(0.0715)	(0.412)	(0.0979)	(0.277)
GAS	0.0164	-0.0318	-0.278	0.0227	-0.142
	(0.0107)	(0.0758)	(0.395)	(0.102)	(0.254)
Constant	1.034***	2.413***	1.578**	0.0778	8.289***
	(0.0293)	(0.184)	(0.705)	(0.249)	(0.728)
Observations	739	670	760	760	666

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 25: Outcome Family 2

	(1) Comparison w/ Other Districts	(2) Corruption Impedes Dev	(3) Project Transparency
CSO	0.0300	0.0344	0.0214
	(0.103)	(0.0895)	(0.0878)
GAS	0.139	-0.00193	0.0481
	(0.0931)	(0.105)	(0.0774)
Constant	-0.481***	0.282*	0.411**
	(0.147)	(0.158)	(0.195)
Observations	760	760	760

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 26: Outcome Family 3

	(1) Participation Improves Dev	(2) Time w/ Constituents	(3) Need-Based Location
CSO	0.296***	2.718*	-0.0395
	(0.104)	(1.471)	(0.0405)
GAS	-0.0173	0.482	0.0405
	(0.115)	(1.257)	(0.0424)
Constant	4.191***	37.30***	0.875***
	(0.296)	(8.381)	(0.174)
Observations	760	760	732

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 27: GSAM Indicators

	(1) Heard of GSAM	(2) Citizens Mention GSAM	(3) Gov Members Mention GSAM	(4) District Responded to GSAM	(5) Citizen Involvement Changed
CSO	0.366***	0.360***	0.368***	0.477***	0.606***
	(0.0355)	(0.0554)	(0.0551)	(0.0665)	(0.143)
GAS	0.311***	0.201***	0.272***	0.324***	0.420***
	(0.0358)	(0.0582)	(0.0598)	(0.0649)	(0.150)
Constant	0.540***	0.0842	0.230**	0.537***	4.009***
	(0.0774)	(0.0847)	(0.0882)	(0.0934)	(0.197)
Observations	755	480	489	353	314

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Annex 5.3: Administrator Outcomes: Secondary Indicators

Table 28: Secondary Indicators

	(1) Political Influence	(2) Oversight Problems	(3) Win Votes w/ Projects	(4) Split Projects	(5) Contracting Standards	(6) DA Budget Changes After Passage
CSO	0.00898	-0.0854	0.0891	0.00574	0.123	-0.0119
	(0.0148)	(0.0844)	(0.114)	(0.0352)	(0.0913)	(0.0802)
GAS	-0.00823	0.0204	0.0329	-0.0297	0.0394	0.0696
	(0.0142)	(0.0892)	(0.123)	(0.0371)	(0.0837)	(0.0824)
Constant	0.0664	2.859***	5.139***	0.608***	0.104	3.647***
	(0.0531)	(0.298)	(0.229)	(0.107)	(0.312)	(0.329)
Observations	748	760	760	760	760	760

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

ANNEX VI: POLITICIAN OUTCOMES

Annex 6.1: Main Results

Figure 28: Outcome 1, Politicians

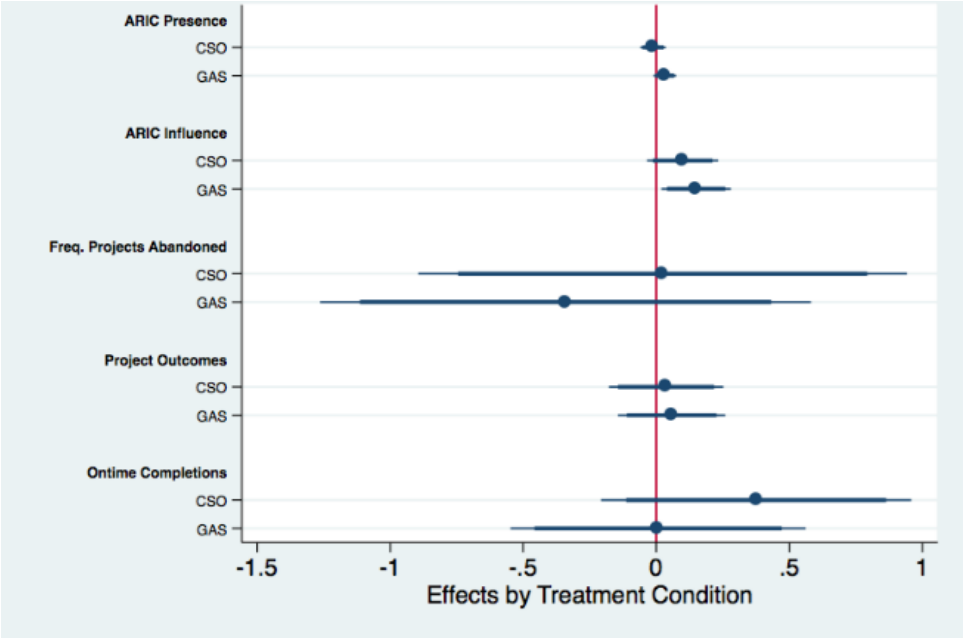


Table 29: Outcome Family 1

	(1) ARIC Presence	(2) ARIC Influence	(3) Freq. Projects Abandoned	(4) Project Outcomes	(5) Ontime Completions
CSO	-0.0121	0.0994	0.0239	0.0371	0.375
	(0.0240)	(0.0679)	(0.464)	(0.109)	(0.295)
GAS	0.0322	0.150**	-0.341	0.0578	0.00690
	(0.0214)	(0.0660)	(0.466)	(0.102)	(0.281)
Constant	1.025***	2.362***	2.870***	-0.253	7.005***
	(0.0549)	(0.197)	(0.583)	(0.297)	(0.644)
Observations	736	536	736	736	603

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Figure 29: Outcome 2, Politicians

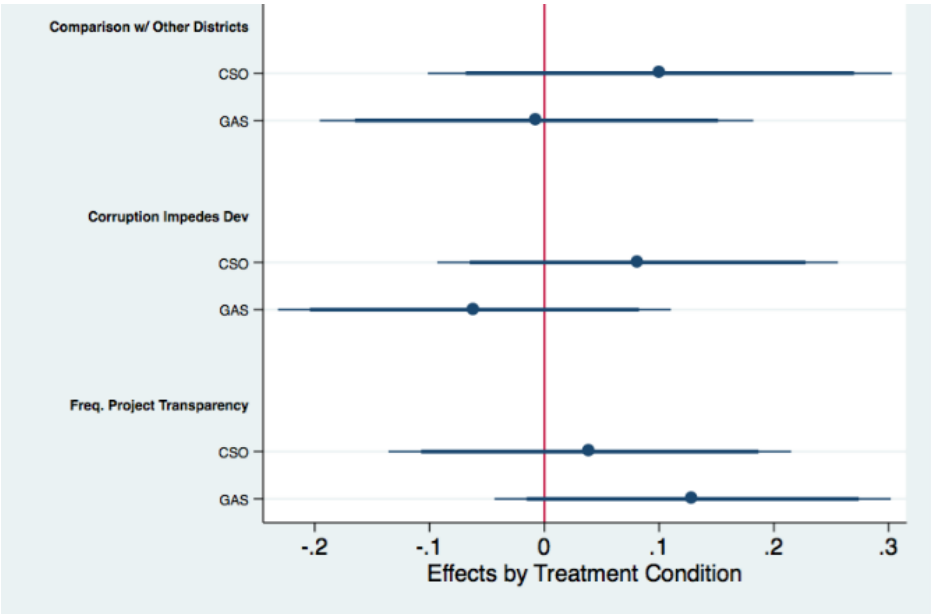


Table 30: Outcome Family 2

	(1) Comparison w/ Other Districts	(2) Corruption Impedes Dev	(3) Project Transparency
CSO	0.101	0.0812	0.0396
	(0.102)	(0.0883)	(0.0887)
GAS	-0.00680	-0.0609	0.129
	(0.0954)	(0.0866)	(0.0873)
Constant	0.0916	0.363	-0.407
	(0.137)	(0.409)	(0.498)
Observations	736	736	736

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

Figure 30: Outcome 3, Politicians

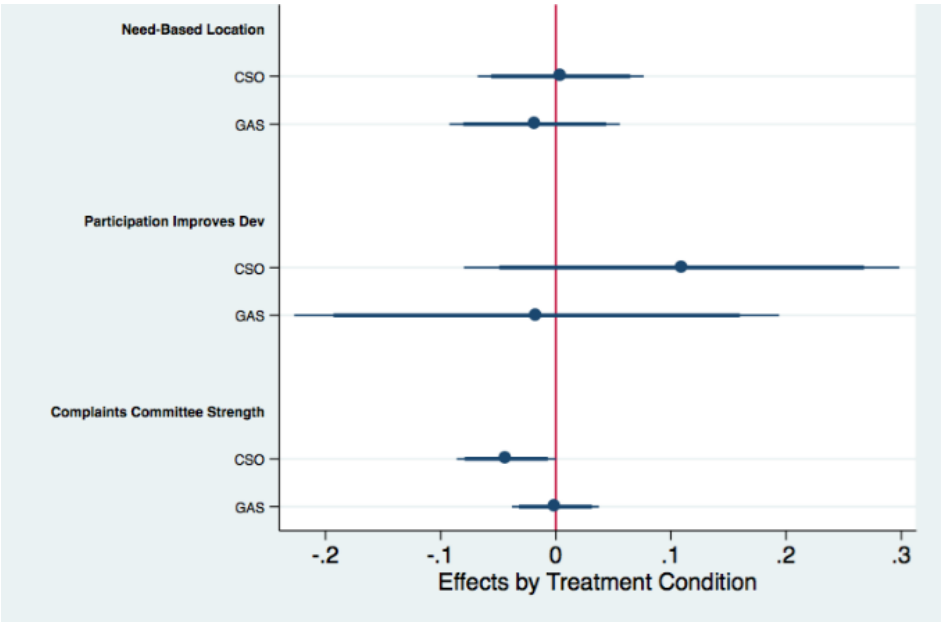


Table 31: Outcome Family 3

	(1) Participation Improves Dev	(2) Time w/ Constituents	(3) Need-Based Location
CSO	0.109	-1.966	0.00419
	(0.0956)	(1.540)	(0.0365)
GAS	-0.0166	0.00180	-0.0183
	(0.107)	(1.630)	(0.0375)
Constant	4.472***	22.43***	0.943***
	(0.238)	(4.237)	(0.0623)
Observations	736	736	720

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Figure 31: GSAM Awareness and Engagement, Politicians

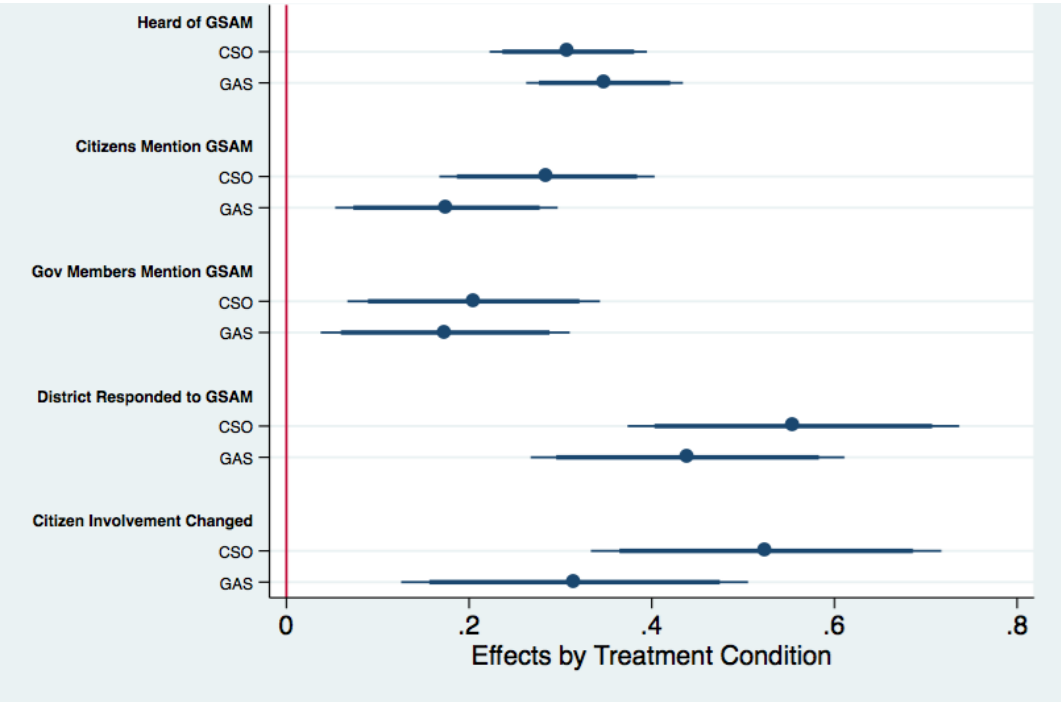


Table 32: GSAM Indicators

	(1) Heard of GSAM	(2) Citizens Mention GSAM	(3) Gov Members Mention GSAM	(4) District Responded to GSAM	(5) Citizen Involvement Changed
CSO	0.309***	0.285***	0.205***	0.555***	0.525***
	(0.0436)	(0.0596)	(0.0699)	(0.0916)	(0.0969)
GAS	0.348***	0.175***	0.174**	0.439***	0.316***
	(0.0434)	(0.0616)	(0.0690)	(0.0867)	(0.0959)
Constant	0.398**	-0.101	0.214***	0.401***	2.797***
	(0.171)	(0.109)	(0.0807)	(0.0848)	(0.117)
Observations	722	401	404	244	281

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Annex 6.2: Politician Outcomes with Controls

Table 33: Outcome Family 1

	(1) ARIC Presence	(2) ARIC Influence	(3) Freq. Projects Abandoned	(4) Project Outcomes	(5) Ontime Completions
CSO	-0.0144	0.0872	0.170	0.0363	0.311
	(0.0240)	(0.0635)	(0.474)	(0.106)	(0.299)
GAS	0.0355	0.0945	-0.165	0.0714	-0.0281
	(0.0221)	(0.0654)	(0.477)	(0.106)	(0.300)
Constant	1.056***	2.271***	1.948**	-0.455	6.537***
	(0.0622)	(0.256)	(0.949)	(0.314)	(0.824)
Observations	736	536	736	736	603

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 34: Outcome Family 2

	(1) Comparison w/ Other Districts	(2) Corruption Impedes Dev	(3) Project Transparency
CSO	0.153	0.130	0.0104
	(0.0968)	(0.0817)	(0.0887)
GAS	0.0261	-0.0589	0.106
	(0.0904)	(0.0809)	(0.0866)
Constant	-0.123	0.190	-0.592
	(0.209)	(0.401)	(0.513)
Observatic	736	736	736

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 35: Outcome Family 3

	(1) Participation Improves Dev	(2) Time w/ Constituents	(3) Need-Based Location
CSO	0.106	-2.051	0.00843
	(0.0987)	(1.505)	(0.0376)
GAS	-0.0490	-0.142	-0.0162
	(0.109)	(1.563)	(0.0394)
Constant	4.535***	21.72***	0.960***
	(0.302)	(5.366)	(0.0869)
Observations	736	736	720

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Table 36: GSAM Indicators

	(1) Heard of GSAM	(2) Citizens Mention GSAM	(3) Gov Members Mention GSAM	(4) District Responded to GSAM	(5) Citizen Involvement Changed
CSO	0.297***	0.322***	0.251***	0.549***	0.674***
	(0.0441)	(0.0616)	(0.0747)	(0.102)	(0.101)
GAS	0.329***	0.180***	0.208***	0.420***	0.431***
	(0.0425)	(0.0634)	(0.0722)	(0.0948)	(0.110)
Constant	0.489**	-0.0548	0.274*	0.512**	2.542***
	(0.222)	(0.166)	(0.161)	(0.242)	(0.279)
Observations	722	401	404	244	281

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

Annex 6.3 Politician Outcomes: Secondary Indicators

Table 37: Secondary Indicators

	(1) Oversight Problems	(2) Win Votes w/ Projects	(3) Split Projects	(4) Contracting Standards	(5) DA Budget Changes After Passage	(6) Electoral Loss from Corruption
CSO	0.0196	-0.0289	-0.0293	0.0383	0.0307	5.248
	(0.0958)	(0.0970)	(0.0431)	(0.0952)	(0.0844)	(4.638)
GAS	0.0864	0.0869	0.00354	0.0941	0.103	7.384
	(0.0957)	(0.0947)	(0.0447)	(0.0966)	(0.0870)	(4.803)
Constant	5.686***	4.748***	0.469**	0.0785	3.584***	16.40
	(0.114)	(0.114)	(0.222)	(0.413)	(0.240)	(12.93)
Observations	736	736	736	736	736	736

Standard errors in parentheses
* p<.1 ** p<.05 *** p<.01

ANNEX VII: DISTRICT-LEVEL OUTCOMES

Figure 32: District Outcomes: Project Completion

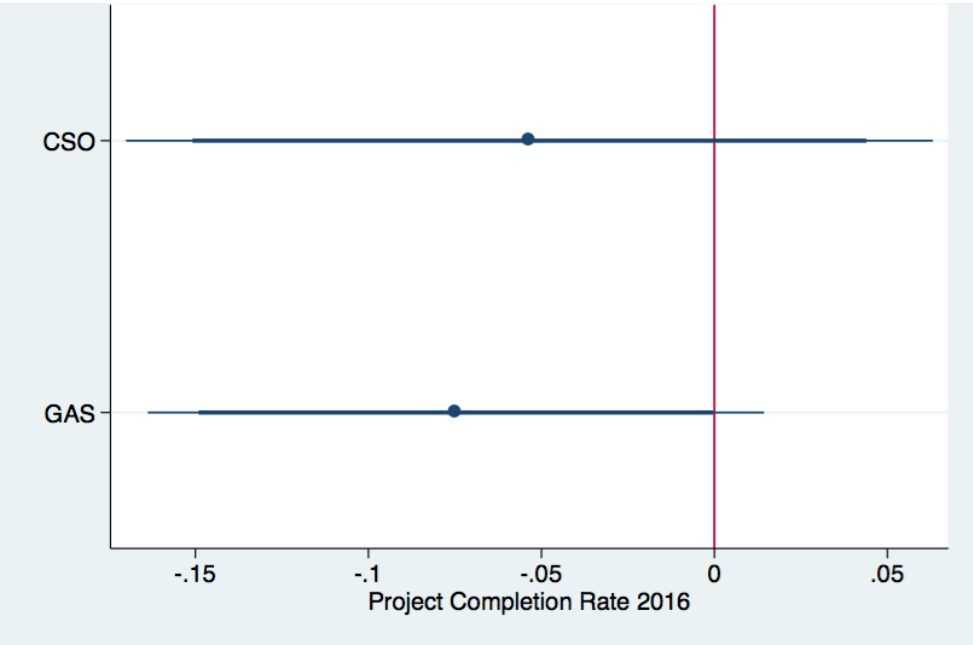


Figure 33: District Outcomes: Irregularities, Normalized by Project Count

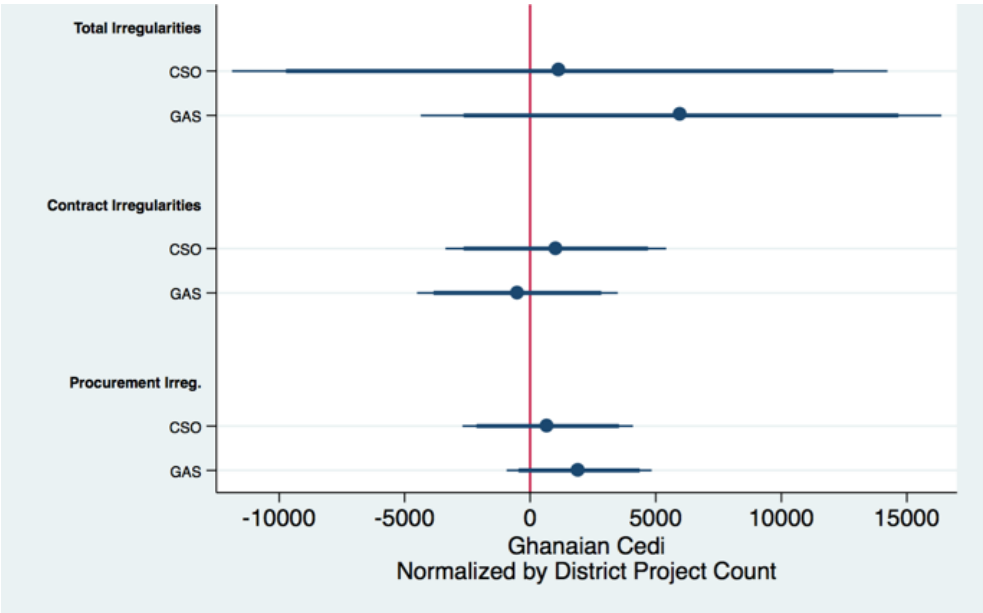


Table 38: District Outcomes (Controlling for Number of District Projects)

	(1) Completion Rate	(2) Contract Irregularities	(3) Procurement Irregularities	(4) Total Irregularities
CSO	-0.0535	21962.7	14228.5	73630.9
	(0.0583)	(42703.3)	(26856.7)	(84734.1)
GAS	-0.0747*	-32367.4	6873.1	-17547.0
	(0.0445)	(39056.7)	(19466.8)	(68624.1)
Constant	0.290*	30108.3	57703.9	187163.0
	(0.169)	(48076.0)	(48636.0)	(116355.2)
Observatic	114	114	114	114

Standard errors in parentheses

* p<.1 ** p<.05 *** p<.01

U.S. Agency for International Development
1300 Pennsylvania Avenue NW
Washington, D.C. 20004