Local Media and Government Responsiveness:
Evidence from Tanzania

Dylan W. Groves [*]

November 17, 2022

Abstract

Does local journalism improve government responsiveness? Independent journalism is considered a cornerstone of accountable governance, but its impact is unproven outside of developed democracies. I argue local journalism improves service delivery in developing countries by helping central governments monitor district-level bureaucrats. I test the argument through a national-scale randomized controlled trial of local investigative journalism in Tanzania. In collaboration with 15 regional radio stations, I identified 206 communities suffering from service delivery problems like flooded roads, broken water points, absent teachers, and missing medical supplies. I then randomly assigned 103 communities to the treatment group and 103 communities to a pure control condition. Journalists from local radio stations visited treatment communities and investigated the source of the service delivery problem, broadcast 20 minute news stories about their findings on regional radio, and conducted follow up investigations several months later. Independent auditors visited all communities 7 months after the investigations were broadcast and scored changes in the targeted service delivery problem. I find that treated communities received higher audit scores on average (coefficient = 0.25 standard deviations, randomization inference p-value = 0.026), amounting to one road or water point repair in every four treated communities. Investigative reports by local journalists spurred responses by un-elected government ministries but *not* citizens, local bureaucrats, or members of parliament. This study offers the first experimental evidence that local journalism improves government performance outside of a developed democracy.

1   Introduction

Citizens in many low-income countries lack access basic public goods and services like safe drinking water, traversable roads, and adequate healthcare and education. A fundamental question in political science is which institutions tighten the link between what citizens need and what governments provide (Przeworski et al. 1999). In theory, the spotlight of local news media offers a promising tool for ensuring responsive governance. Classic models of electoral accountability and a growing body of empirical research support the contention that press coverage improves government performance in developed democracies by allowing citizens to select more effective politicians and to motivate those politicians once they are in office (Fearon 1999; Barro 1973; Besley and Burgess 2002; Banerjee et al. 2020; Snyder and Strömberg 2010).

But the promise of independent media is less certain outside of developed democracies. Evidence for the effect of transparency campaigns in developing countries is decidedly mixed, and it is especially sparse for tangible outcomes like the provision of public services (Raffler 2019; Grossman and Michelitch 2018; Kosack and Fung 2014; Malesky et al. 2012; Anderson et al. 2019). Meanwhile, an experimental assessment of local journalism has, to my knowledge, never been conducted.

This paper uses a national-scale field experiment in Tanzania to investigate the influence of local journalism investigations on the provision of public services. I worked with radio stations and journalists in 15 regions to identify service delivery failures in 206 communities, each serving approximately 1,500 citizens. I randomly selected 103 communities to be the target of local media investigations which were collectively entitled *Wakati Ni Sasa* (*WNS*) (*The Time is Now*).[1] The remaining 103 communities made up a control condition. Investigations focused on one problem in each community specified prior to randomization, such as a broken bridge, unpotable water, absentee nurses or an overcrowded school. In treatment communities, journalists spoke with citizens, elected officials, local bureaucrats and relevant ministries to document the problem, then produced and broadcast 20 minute radio news programs about their findings. Journalists returned to treatment communities 4-6 months later for follow-up reports. Survey evidence confirms that *WNS* investigations were widely remembered by government officials 7 months after they aired, and the reports generated an expectation that future service delivery failures would also receive media attention.[2]

To measure the effects of the *Wakati Ni Sasa* investigations, a team of 16 enumerators conducted audit surveys of all communities in the sample 7 months after the initial investigations were broadcast. In each community, auditors conducted 2-3 surveys with local leaders and bureaucrats, collected photographic, video, and audio documentation of service delivery changes, and compiled 1-2 page qualitative assessments of how the targeted problem was addressed

---

[1] Audio and short summaries of *Wakati Ni Sasa* stories can be found here

[2] 75% of government officials remembered a local media investigation in their village about the pre-specified problem, 38 percentage points higher than the control group (RI *p*-value < 0.001. 62% of village chairpersons said a future report was "very likely", 15 percentage points higher than the control group (RI *p*-value = 0.005)

over the project period. Auditors then filled out report cards by assigning a score betwen -10 and 10 to overall improvements in the targeted public good. They also assigned grades to the responses of citizens, non-governmental organizations, and different government actors. To test the reliability of the auditor's scores, an independent auditor, blind to the purpose of the study, filled out another report card using only documented evidence collected by the original auditor.

The primary finding is that *Wakati Ni Sasa* reports improved public service provision in Tanzania. Auditors gave better ratings to public service improvements in treatment communities than in control communities 7 months after the initial *WNS* reports were aired (coefficient = 0.25 standard deviations, randomization inference p-value = 0.026). The effects correspond to one repaired road or water point for every four villages where a *WNS* investigation was conducted. The effect of *WNS* reports was driven by actions taken by governments rather than citizens or non-governmental organizations.

The second finding is that local press coverage stimulated responses from unelected ministry officials but not village-level government officials or members of parliament. To understand the reasons for this finding, I draw on more than 12 months of field research shadowing journalists and radio stations between 2019 and 2020 and 50 interviews with journalists, politicians, and ministry officials. I find that local media investigations improve public service provisions by helping the ruling party monitor district-level ministry officials. These officials respond to *WNS* investigations because negative news coverage about the ruling party threatens their career prospects within the bureaucracy. I leverage plausibly exogenous variation in top-down sanctioning capacity from the rotation of Regional Commissioners to show that the influence of *WNS* investigations on district bureaucrats was strongest where threats of central government sanctions were most credible.

There are important limitations to these findings. The experiment was implemented outside of an election year and after annual budget allocations, when mechanisms of electoral accountability are less likely to be operative, and the experiment took place only over the course of 7 months, so the full range of plausible citizen and government responses may not have had time to manifest. It is therefore possible that different channels of media influence may be more salient under different experimental conditions. In-person interviews with citizens and some leaders was made impossible by COVID-19 restrictions, so outcomes are limited to audit reports. Finally, the experimental design cannot rigorously determine whether media reports influenced the *amount* of resources provided to citizens or simply their *allocation*. Each of these concerns lay groundwork for future research.

This paper contributes to literature on independent media, hegemonic party regimes, and state-building. The first contribution is to demonstrate that local press coverage influences the distribution of public goods and services even in the "hard-case" of (i) a low-income country (ii) ruled by a hegemonic party regime (iii) outside of an election year. The rapid growth of independent media in developing countries has engendered a robust literature on media's influence on a range of social, economic and political outcomes (Durante et al. 2019). However, few studies have

investigated the impact of local media stations performing one of their most basic functions: reporting on the everyday challenges facing the communities they cover. This paper builds on a small but growing literature showing that problem-focused news coverage (sometimes referred to as "development journalism") represents a promising tool for promoting good governance even outside of industrialized democracies (Sen 2001).

The paper also provides insights into processes of political accountability and distribution in hegemonic party regimes that predominate in Sub-Saharan Africa (Levitsky and Way 2012). Popular models of government responsiveness emphasize accountability relationships between citizens and politicians (Grossman et al. 2020), citizens and local bureaucrats (Kosack and Fung 2014; Björkman and Svensson 2009), and politicians and bureaucrats (Raffler 2019; Slough 2020). Without discounting these accountability relationships, this paper highlights how media facilitates top-down accountability *within* developing country bureaucracies (Lorentzen 2014).

More broadly, the paper contributes to our understanding of state-building outside of industrialized democracies. Classic theories of state development posit that state capacity depends upon a government's ability to gather information about its citizens and bureaucracy, and recent scholarship has linked this "informational capacity" to a range of positive economic and social outcomes (Bowles 2022; Brambor et al. 2020; Lee and Zhang 2017). Following Egorov et al. (2009) and Lorentzen (2014), the results of this paper suggest that media plays an important role building the informational capacity of developing states.[3]

Finally, this paper has important implications for policymakers. Global democratic retrenchment in the last two decades has raised important questions about which institutions can bolster government responsiveness in the absence of a strong opposition party. At the same time, recent research on misinformation and media capture has raised questions about whether mass media helps or hinders the prospects for democratic accountability in the developing world (Guriev and Treisman 2019). The results presented here suggest that investing in local news media offers a cost-effective avenue for promoting economic development that reinforces, rather than circumvents, institutions of responsive governance.

This essay is organized as follows. I begin by outlining a top-down model of media influence drawn from more than 12 months shadowing investigative journalists and radio stations. I then briefly describe the politics, media, and service delivery in the Tanzanian context. Having set the backdrop for the study, I describe the *WNS* treatment and lay out the key features of the experimental design: the selection of radio stations, sampling of communities, and random assignment, the measurement of outcomes, and the estimation procedures. I then assess the effects of the treatment on public goods provision and responses by specific actors, and explore both quantitative and qualitative evidence for mechanisms. I conclude by discussing the implications of the findings and suggesting avenues for further investigation.

---

[3]The origins of mass media are deeply tied to processes of state development: the earliest newspapers developed to meet the demand of European political elites for news about periphery (Pettegree 2014).

2 Theory: Journalism and Responsiveness in Hegemonic Party Regimes

How might media influence government responsiveness in the absence of electoral pressures generated by a strong opposition party? I argue that media investigations help hegemonic party regimes overcome a central obstacle to effective governance: monitoring the performance of district-level bureaucrats. Central governments in hegemonic party regimes face a classic principal-agent problem in which they would like district-level ministry officials to deliver services and respond to citizen needs, but can only imperfectly monitor those ministry officials' performance (Lorentzen 2014; Egorov et al. 2009). District-level ministry officials would prefer to shirk and expropriate funds, but can be spurred to execute the central government's priorities if they are sufficiently worried that the central government would observe and sanction incompetence and malfeasance. When district-level ministry officials believe that a local service delivery failure is likely to be investigated by the media, they become more motivated to address the problem (Anderson et al. 2019).

In Tanzania, government resources for services like water, health care, education, and roads are distributed at the district level by local government authorities (LGAs) or district-level ministry offices (see Figure 2). Tanzania's President appoints Regional Commissioners to oversee the execution of these projects (Mdee and Mushi 2021). Regional Commissioners are often referred to as the "eyes and ears" of the President at the district and regional level, and they have the authority to influence promotions and demotions within the regional bureaucracy. District-level ministry officials are highly motivated to curry favor (and avoid falling out) with Regional Commissioners in order to advance their careers within the ruling regime (Fukuyama 2014; Cheeseman et al. 2021). However, the oversight capacity of Regional Commissioners is imperfect, creating incentives to shirk for district-level bureaucrats (Mdee and Mushi 2021; Carlitz 2017).

In this context, local media investigations can serve an important function for the Tanzanian ruling party by exposing governance failures at community level. According to an award-winning local journalist in Mbeya,

> "*The (district-level) ministry officials have to worry that if they don't do anything, they will get in trouble with the government because they have a budget so as to create those roads, so if a journalist goes and asks about that budget, the ministry officials have a worry, usually about the regional commissioner or of course from the Minister.*"

This role for media is deeply rooted in Tanzania's political history. Veteran journalists from Tanzania's era of single-party politics recall waiting by the phone after publishing investigations into local government malfeasance in the hopes that Tanzanian's first president, Julius Nyerere, would call them and thank them for alerting him to the issue (Mwafissi, interview). Media's role facilitating top-down accountability is also reminiscent of descriptions of media influence in China (Pan and Chen 2018; Anderson et al. 2019; Truex 2016). However, media's influence on top-down channels of accountability in a hegemonic party regime has yet to be tested experimentally.

3   Tanzanian Context

Three factors make Tanzania an apt context for studying media and governance in a hegemonic party regime. First, Tanzania's ruling party, Chama Cha Mapinduzi (CCM), exhibits a level of hegemonic control and institutionalization that reflects many shared features of hegemonic party regimes in Sub-Saharan Africa and throughout the developing world. Second, Tanzania's media industry is underdeveloped but growing rapidly, so Tanzania's citizens and government officials are familiar with local press reports without being saturated by them. Finally, public service delivery in Tanzania is limited, leaving the door open for external interventions to improve development outcomes (Croke 2021).

3.1   Dominant Party System

Tanzania has been governed by a single political party, Chama Cha Mapinduzi (CCM), since its unification in 1964.[4] Under Presidents Benjamin Mkapa (1995-2005) and especially Jakaya Kikwete (2005-2015), the space for political dissent widened significantly. However, after a closer-than-expected national election in 2015, newly-elected president John Magufuli enacted several measures measures that critics argued limited media freedom and internal dissent. Most opposition leaders boycotted local elections in 2019 and the 2020 (Cheeseman et al. 2021). As a result, at the time of the *WNS* intervention in 2021, CCM controlled all local elected offices and 97% of seats in parliament. President Magufuli passed away in April 2021, just as the *Wakati Ni Sasa* intervention was being planned. His position was filled by Samia Hassan, who promised to ease restrictions on opposition organizing and media freedoms, although the extent of this policy transformation remains unclear (Cheeseman et al. 2021).

Through this history, CCM has rarely faced a meaningful threat from opposition parties. CCM has never won less than 58% of the presidential vote or 70% of seats in parliament (Morse 2019, 2012; Weghorst 2022). CCM's electoral hegemony is attributable in part to a sophisticated and highly institutionalized party apparatus with operations that extend to the village level.[5] CCM is broadly representative of other hegemonic party regimes (Morse 2019; Weghorst 2022)

3.2   Growth of Independent Mass Media

Like citizens of many post-colonial states, Tanzanians have witnessed a dramatic expansion of local and independent media over the last thirty years (Sturmer 1998). From independence in 1962 until 1993, Tanzanians were able to access just two radio channels: state-owned Radio Tanzania Dar es Salaam and Sauti ya Injilli (Voice of the Gospel), a non-news radio station owned by the Catholic Church which was introduced in the 1980's (Sturmer 1998). After political liberalization and the repeal of Tanzania's most onerous media restrictions in 1993, the number of independent radio stations climbed to three in 1995 and 11 in 2005. After 2005, the number of independent mass media

---

[4]Prior to unification, Tanzania was ruled by CCM's precursor, TANU, from independence in 1961 through 1964. It was a one party state until the introduction of multipartyism in 1992.

[5]Often called the "ten cent system"

houses grew dramatically, from 25 in 2005 to 125 in 2015 (see Figure 1). Today, there are just over 200 radio stations.

Figure 1: **Growth of Independent Radios in Tanzania**
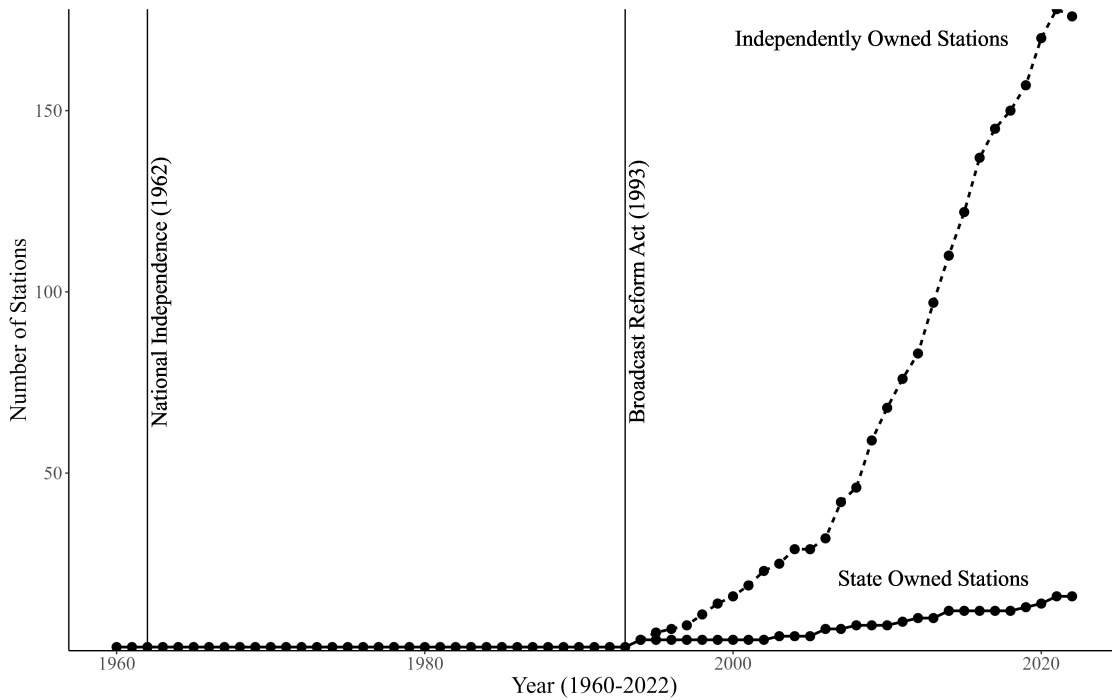
While many of Tanzania's newly minted media houses are focused on urban centers like Mwanza and Dar es Salaam, about 70% of Tanzania's registered radio stations are based in regions that, prior to 2005, had never carried a regionally-focused news outlet. These local and independent radio stations are often owned by local business people and politicians (47%), religious organizations (21%), and community-based organizations (30%). While most privately-owned radio stations focus on music and entertainment, about one quarter are focused on news, politics, and social issues, and almost all local radio stations broadcast daily local news coverage and employ teams of journalists (Katunzi and Spurk 2019; Spurk and Dingerkus 2017). In comparison to other East African countries, however, Tanzanian journalists have less training and fewer resources (Katunzi and Spurk 2019). Nonetheless, local media investigations are a standard component of local radio programming in Tanzania. The Tanzanian Development Information Organization (TADIO), an umbrella organization of more than 34 local radio stations representing every region of Tanzania, shares weekly investigative news reports broadcast by its partners. These stories range from sharing a community's grievances about the the distances children walk to secondary school to tracking government responses to water shortages.[6]

In sum, Tanzania's media sector is growing rapidly but remains underdeveloped relative to other countries in East Africa. Most citizens can listen to or read press coverage about social and political issues, but it is rare that

---

[6]https://radiotadio.co.tz/dodomafm/2022/09/01/umbali-wa-shule-wapelekea-baadhi-ya-wanafunzi-pandambili-kuacha-shule/,
https://radiotadio.co.tz/nurufm/2022/09/05/mtendaji-na-katibu-bandia-mbaroni-kwa-uchochezi-mkoani-morogoro/

their local grievances will be the subject of press reports, especially in rural regions. As I discuss in more detail below, the *Wakati Ni Sasa* press reports discussed in this paper were not significantly different than the standard practice of Tanzania's local radio stations. Instead, *WNS* influenced the frequency, targeting, and quality of local media investigations.

3.3   Delivery of Public Goods and Services

Tanzanian citizens often lack access to basic public goods and services like health care, education, water, and roads, a bundle of basic public goods often collectively referred to as "maendeleo" (development) by Tanzanian citizens. In Afrobarometer surveys, just 21% of Tanzanian citizens reported having access to piped water, 68% to schools, and 37% to health services in their community. More than half of low-income Tanzanians report moderate or high difficulty gaining access to medical care or water, sanitation, and electricity services.

Since 1999, the Tanzanian government has pursued a strategy of "decentralization by devolution" (Mclellan 2021; Mdee and Mushi 2021). The centerpiece of this strategy has been transferring authority for implementing development programs from the central government to officials at the district level. In theory, decentralization brings development decisions to the level of the community, improving targeting, responsiveness, and citizen input. In practice, decentralization has been both less effective and less complete than originally planned (Mdee and Mushi 2021). Responsibility for water and roads projects was first devolved then re-centralized in 2016, while responsibility for health and education remains blurred. The result is that citizens are often unaware which government bodies are responsible for guaranteeing the provision of different government services.

During in-depth interviews, village, ward, and district officials outlined three routes to government service delivery (see Figure 2). In the "standard route" to public goods provision, the parliament allocates budgetary support to districts (known as Local Government Authorities, or LGAs) earmarked for specific development priorities. District executive officers work with elected district councils to allocate funds to wards and villages. This route is particularly common for the provision of health and education services. A second route runs through national ministries and parastatals, such as the Rural Water Supply and Sanitation Agency (RUWASA), the Tanzania Rural and Urban Roads Agency (TARUA), and the Tanzania Electric Supply Company Limited (TANESCO), which maintain offices in each region and are empowered to directly pursue development projects at the village and ward level. When funding from district councils and ministries is not forthcoming, village and ward governments often raise money directly specific development projects by taxing local businesses or soliciting citizen contributions. Finally, in some cases Members of Parliament contribute money and supplies to ongoing development projects, especially during election years. In the absence of government support, public goods and services are often provided by citizen self-help initiatives and non-governmental organizations.
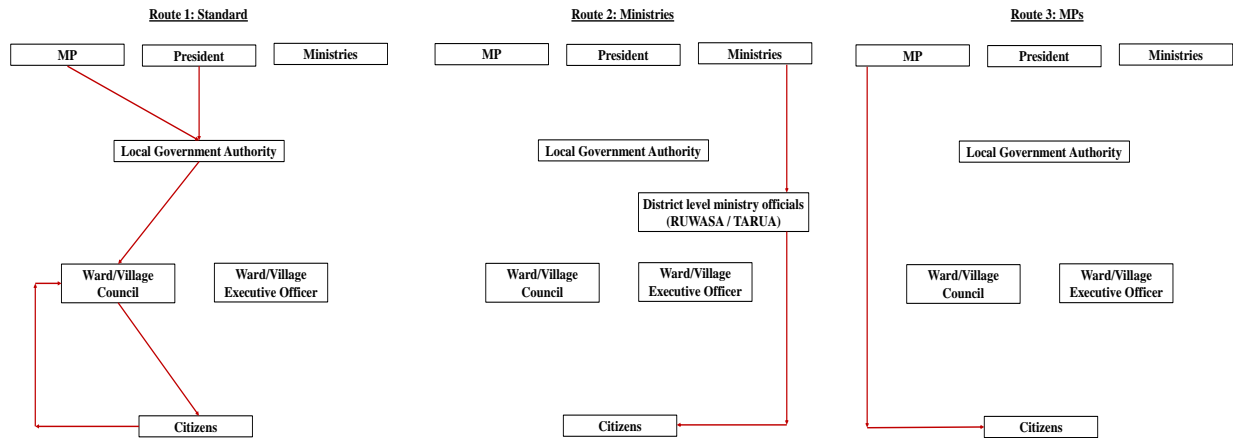
Figure 2: Three routes to service delivery

4 Intervention - *Wakati Ni Sasa*

The focus of this paper is the influence of local investigative journalism on the provision of public goods like water, health care, education, and roads. I spent more than 12 months shadowing local journalists in an effort to develop a template for local news reporting that reflected standard procedures among high quality reporters. I then collaborated with radio station partners and journalists standardize a template for reporting, producing, and broadcasting local investigations about service delivery failures in a project entitled *Wakati Ni Sasa* (*The Time is Now*).[7] The *WNS* intervention represents a bundle of three related components: (1) investigations into the causes and consequences of service delivery problems; (2) production and broadcast of 20 minute news reports about the results of the investigations; and (3) follow-up contacts with responsible officials to track responses to the initial investigations. The procedure represents "best practices" for trained journalists in Tanzania (Katunzi and Spurk 2019). Figure 3 shows the timeline of piloting, reporting, and data collection in the study.

Figure 3: **Study Timeline**
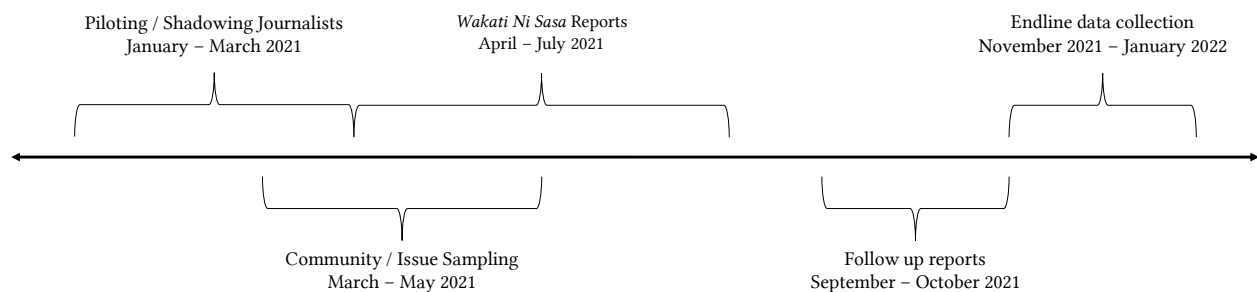
Note: Data on radio station growth and ownership comes from the Tanzanian Communications Regulatory Authority (TCRA).

---

[7] Audio recordings and written summaries of reports are available here.

Wakati Ni Sasa journalists began each investigation by reaching out to local leaders and key informants in targeted communities. Preliminary outreach served both a pragmatic and ethical function. Pragmatically, preliminary conversations gave journalists a preview of the service delivery problem and let them know whether any background interviews (for example, with water or road engineers) were necessary prior to visiting the community. Ethically, preliminary conversations with government officials reduced the risk of unexpected confrontations with local leaders during the community visit and helped journalists identify safety issues (COVID, crime, or unsafe transportation) they needed to consider before departing.

After conducting outreach, *WNS* journalists travelled to the targeted community to investigate the pre-specified service delivery problem. The journalists were trained to begin their investigations by interviewing a range of citizens about the history of the problem and the way the problem was affecting their community. Journalists then conducted follow-up interviews with government officials. While *WNS* journalists were given wide latitude to pursue their investigations as they and their station manager saw fit, they were expected at minimum to interview the village/street chairperson and executive officer, ward councilor and executive officer, District commissioner, and representative from the relevant ministry. *WNS* were also required to contact the relevant Member of Parliament to request comment. Requiring *WNS* journalists to contact a range of government officials adhered to best practices in investigative journalism[8] as well as Tanzanian media law.

Community visits usually lasted between one and three days, and follow-up interviews with higher level government officials lasted anywhere between one week and several months. The most common reason for extended reporting times was the refusal of District and Regional officials to make themselves available for interviews. *WNS* partners then produced 15 minute narrative reports for radio broadcast. By agreement, the reports always consisted of an introduction explaining the broader relevance of service delivery issue, interviews with citizens emphasizing how the service delivery issue was affecting their lives, and interviews with relevant government officials about the history of the problem, the reasons for government inaction, and the government's plans for addressing the issue in the future.

All reports were broadcast twice over the course of a week, and listeners were given the opportunity to call or text their comments. In addition to the radio broadcast, reporters were free to post pictures and discussion about the stories to their website and social media. In anticipation of the broadcast, *WNS* journalists called key citizen informants and government officials to inform them that the episode would be put on air. They also informed all parties that they would conduct a follow-up episode in the ensuing months to track the community's progress towards resolving the issue.

Three to four months after the first broadcasts, journalists conducted follow-up investigations in targeted communities. Follow-up reports were shorter than the original reports, and included phone call interviews with citizens

---

[8]See Appendix for details on journalism training

and government officials. The reports were conducted regardless of whether the government had resolved or failed to resolve the specified issue.

4.1 Was *Wakati Ni Sasa* realistic?

How closely did *Wakati Ni Sasa* investigations reflect real-world news reports in Tanzania? As the description of media context in Tanzania makes clear, local news investigations are a weekly feature of at least one (and often more) radio station in every region in Tanzania. I took several steps to ensure that the *WNS* reporting template reflected standard operating procedures at local media stations. Radio station partners picked journalists for *WNS* who were already on the station's staff, and *WNS* journalists did not receive additional formal training. The sampling process was modelled on the weekly routine radio newsrooms follow to select communities for local reports. The resulting news stories were produced and broadcast entirely by radio station staff. The primary deviation from standard reporting procedures was that journalists submitted the results of their investigations to a research team member, and in some cases received feedback about additional governmental officials they should contact.

Rather than introducing a qualitatively different type of press coverage, then, *WNS* changed the *frequency*, *targeting*, and *quality* of local media investigations. *WNS* changed increased the frequency of media investigations by providing funds to journalists to report and produce more stories than their station managers could otherwise afford. *WNS* influenced the *targeting* of media investigations by introducing a randomized component and by providing transportation funds so that journalists could reach rural villages they otherwise could not reach. Finally *WNS* influenced the *quality* of media investigations at some radio stations by introducing a shared template for reports so that journalists at lower-capacity stations still followed the best practices of high performing journalists I shadowed in the qualitative lead-up to the project.

5 Research Design

The experiment is a matched-pair randomized controlled trial with randomization and outcome measurement at the community level. 103 treatment communities received the *WNS* intervention, which included investigations, broadcasts, and follow-up reports. The remaining 103 communities made up a pure control condition. This section reviews the project's sampling process, random assignment, data collection, and estimation strategy.

5.1 Sampling

5.1.1 Radio Stations

I conducted the experiment in collaboration with 15 local radio stations located across Tanzania. I partnered with radio stations that met three conditions: they were located in mainland Tanzania (rather than Zanzibar), they had district or regional (rather than national) radio broadcast permits from the Tanzania Communications and Regularly Agency (TCRA), and they were identified by professors of journalism at the University of Dar es Salaam and St. Augustine University School of Journalism for broadcasting high quality news reports. I ultimately identified radio

stations in 15 of Tanzania's 26 mainland regions. Excluding Dar es Salaam, regions with selected radio stations are comparable to regions without selected radio stations in average size (35,000 vs 36,000 km^2) and average population (2.0 vs 1.8 million).

5.1.2 Villages/Streets

The study's unit of randomization is the community (referred to as a village in rural districts and a street in urban and peri-urban districts). To identify communities with service delivery problems, radio stations asked listeners to call or send text messages about water, health, education, and infrastructure challenges in their area. They also included these requests after *Wakati Ni Sasa* broadcasts. In the week-long preparation stage, participating radio stations received between 60 and 180 text messages, and continued to receive SMS texts and calls throughout the project period. A sample of texts across different issue areas appears below:

> "*[Ward name] has 4 teachers and does not have a single female teacher and has about 500 students and only 6 classrooms is a big problem problem over the problem. This is [name]*"

> "*I am called [name] of [village] while I have the biggest challenge is the poor school toilets and I really ask the government to help us*

> "*I am called [name] I am found in the village of [village name], the challenges of our village is in the health side, with insurance if you go to another [clinic] you are written to buy [from] the pharmacy, it means the government does not have medicines.*

> "*Water crisis in the Dodoma region of [district name] of [village name] has a shortage of water...Please we need help to solve this problem of water. Water shortage Thank you [phone number]*"

> "*My name is [name] in the village of [village name] Region Dodoma. The challenge facing [school name] children is a road full of water that prevents children from school: [mobile number]*"

Radio stations also learned about service delivery problems through informal networks of informants. Radio stations submitted villages in pairs to facilitate the matched-pair randomization described below. The research team excluded communities if they could not verify the problem through follow-up phone calls to community members, if the proposed investigation posed risks to the safety of citizens or journalists, or if another community in the same ward was already included in the sample. Ultimately, the research team selected 103 village pairs (206 total) for *WNS* program eligibility. This strategy limited the sample to communities where at least some members were aware of the service delivery problem; this study does not consider the role of media in *uncovering* previously unrecognized governance failures.

5.2 Random Assignment

I conducted random assignment to experimental conditions at the community level after grouping communities into matched pairs. Paired communities are generally located in the same district and requested media investigations in the same week. Within each matched pair, one community was randomly assigned to treatment and one

community as randomly assigned to control. Figure 4 shows the geographic distribution of treatment and control communities. Treated communities are colored blue and control communities are colored red, while each shape represents a distinct service delivery issue.

Figure 4: **Random Assignment**
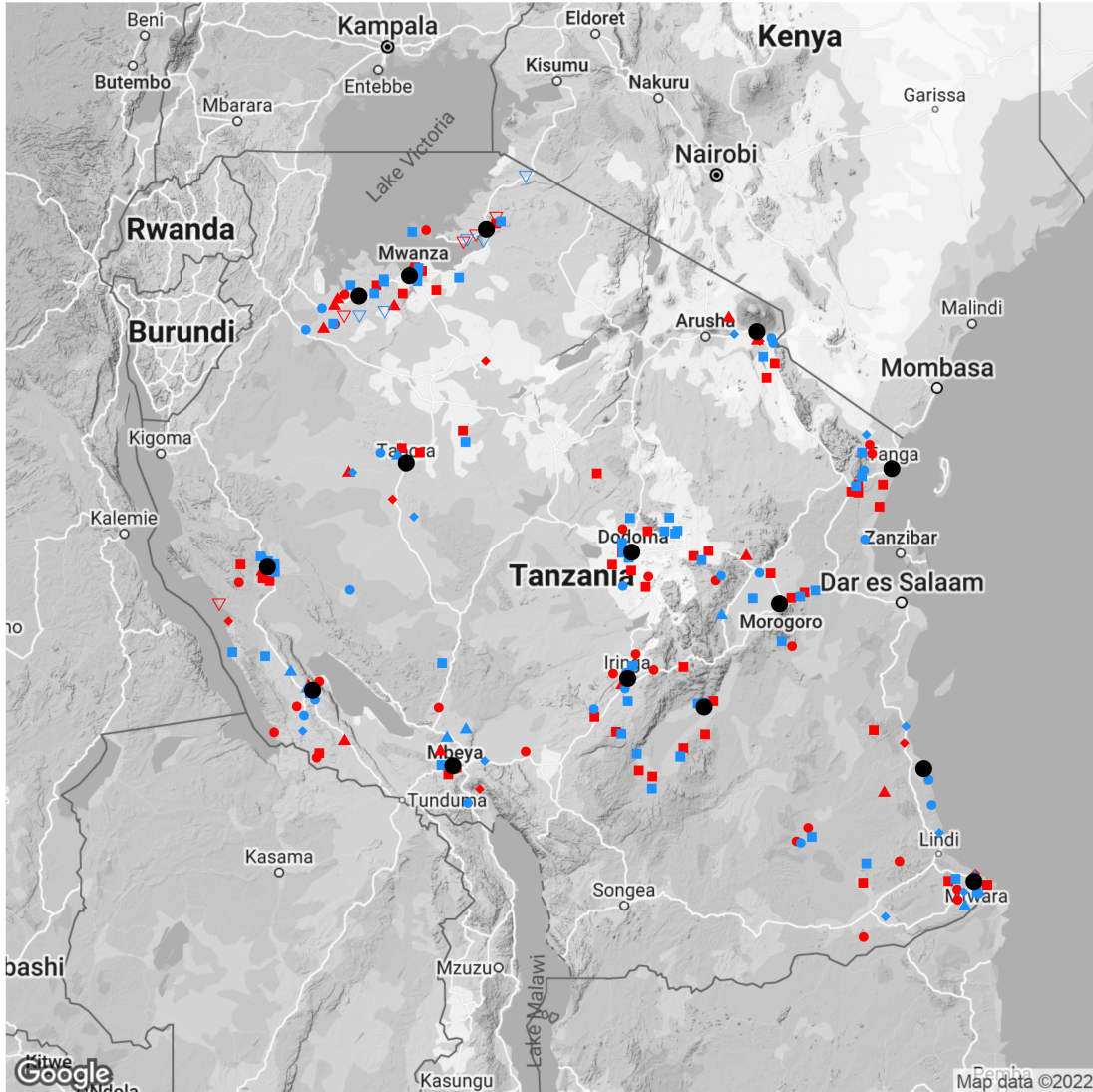
Table 1 demonstrates that random assignment was not significantly correlated with administrative measures of pre-treatment population, electoral competition, or distance from roads, towns, and service centers, nor was it correlated with pre-treatment measures of the nature and severity of the specified service delivery problem.[9] Out of 19 pre-treatment covariates, one (5%) generates a randomization inference p-value of 0.05 or smaller, and only

---

[9]I did not conduct a baseline survey due to concerns that baseline data collection would triggered two potential channels of media influence: mobilizing community focus on a specific service delivery issue and making local leaders feel that their performance was being observed.

two (11%) generate a randomization inference *p*-value of 0.1 or smaller, almost exactly what is expected to occur by chance alone.

Table 1: **Balance on Pre-Treatment Covariates**

Variable	Treatment	Control	RI p-value	Observations
Politics				
Village Established Afer 2015	0.175	0.126	0.370	206
Share CCM Ward Councillor Vote - Ward (2015)	0.611	0.621	0.420	190
CCM Ward Councillor Dummy - Ward (2015)	0.700	0.700	0.850	200
Population				
Male Population	1372.872	1640.483	0.110	165
Female Population	1412.936	1729.782	0.085	165
Household population	583.462	684.954	0.120	165
Service Delivery Problem Severity				
Pre-Treatment Problem Score (0-10)	7.010	6.853	0.580	203
Pre-Treatment Problem Pair Winner	0.059	−0.059	0.455	203
Service Delivery Problem Type				
Problem - Education	0.097	0.087	0.560	206
Problem - Health	0.233	0.262	0.450	206
Problem - Transport	0.107	0.136	0.340	206
Problem - Water	0.515	0.466	0.390	206
Problem - Other	0.049	0.049	0.735	206
Distances (km)				
Distance from Radio Station	44.063	47.526	0.155	206
Distance from Regional Capital	48.836	54.028	0.035	206
Distance from Dodoma	352.708	351.524	0.660	206
Distance from Clinic	33.832	35.842	0.335	206
Distance from Schools	4.067	4.104	0.935	206
Distance from Secondary Roads	21.378	21.815	0.775	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. "Village established after 2015" is a binary variable that takes the value 1 if the village was established after 2015 and 0 otherwise. The remaining politics variables from publicly release vote totals. Dummy variables take the value 1 if the Ward Councillor or Parliamentary seat was occupied by a CCM party member in 2015. Population values are taken from Administrative data shared by Tanzania's Data and Statistics Agency. Data are from 2015, and so is unavailable for newly established villages. Service delivery problem severity data comes from independent audit evidence collected about the original service delivery issue. Service delivery type is a series of binary variables based on the pre-treatment identification of the issue area in each village. Distances are calculated using GPS coordinates of the village and geo-located data on the location of cities and administrative units.

Table 1 also offers a useful profile of the community sample. On average, sampled communities are small and rural: their populations range between 1,000 and 2,000 citizens and are on average 21 kilometers from the nearest paved road and 4 kilometers from the nearest school. Politically, most communities supported the ruling *Chama Cha Mapinduzi* party, but the support was not overwhelming. On average, 61% of voters in sample communities voted for the CCM candidate for Ward Councillor, and the CCM candidate won 70% of 2015 Ward Councillor elections in sampled wards. With respect to the service delivery problems in sampled communities, Table 1 shows a clear focus on water shortages (48% of communities), followed by health services (23%), infrastructure (11%), and education (9%). *WNS*'s emphasis on water service delivery presaged a broader shift in governmental focus on rural water supply in 2022: just as endline data collection was ending, The Tanzanian Ministry of Water and Irrigation announced a country-wide inspection of all government water projects.[10]

---

[10] *WNS* cannot claim credit for Aweso's announcement.

5.3 Outcome Measures

5.3.1 Treatment Uptake

The first group of measurements centers on program implementation. To assess whether *Wakati Ni Sasa* reports were produced and broadcast, I asked radio stations to submit an audio file of the finished report, and assigned a research team member to listen to the station to confirm that the program was broadcast at two agreed-upon dates.

I also assessed whether *Wakati Ni Sasa* reports were noticed and remembered by local leaders in targeted communities. Enumerators conducted interviews with three community leaders: the village chairperson, the village councilor most directly responsible for the targeted service delivery issues (e.g. the head of the education committee), and the chairperson of the largest sub-village. To measure whether local leaders encountered journalists during the investigations, enumerators asked leaders "How many times has a journalist visited this community or street to address the [targeted issue] problem in the past 7 months?" To measure whether leaders heard the reports broadcast on air, enumerators also asked "Have you heard a radio report about a [targeted issue] problem in your community in the past 7 months?" For both responses, I coded the village-level outcome as 1 if any village leader responded "yes" and 0 otherwise. I then constructed an index taking the value 1 if any village leader remembered a journalist investigation about the targeted problem *or* heard a report broadcast about the targeted problem.

I also measured how *WNS* reports influenced leaders' perceptions about the likelihood of future reports. Enumerators asked the village chairperson "How likely do you think it is that a journalist will visit your village/street if it has a [targeted issue] problem in the future?" and coded the response as 1 if the village chairperson said "very likely" and 0 otherwise (very few leaders responded "not at all likely", even in villages that had never received a village report).

5.3.2 Service Delivery

The outcome of primary interest is the quality of core services like education, health care, water provision, and roads. To measure changes in service delivery, trained auditors conducted unannounced audits in every community in the sample 5 to 7 months after the original *WNS* reports aired. Trained auditors, blind to both treatment assignment and the nature of the intervention, contacted located leaders, citizens, and District officials in each community and asked them about their experiences with the pre-specified target problem. Based on these interviews, auditors wrote narrative reports about the original service delivery issue as well as steps that had been taken to address the problem in the previous 7 months. In addition to these qualitative accounts, enumerators collected photographs, videos, and audio recordings to document observable changes in the targeted issue area. They also collected evidence of changes in *other* service delivery areas to test for potential spillover effects. Auditors submitted their findings in the form of 1-2 page audit reports that are available in the Online Appendix.

After collecting interviews and visual evidence, auditors scored service delivery improvements in each commu-

nity. Auditors were asked, "On a score of -10 (much worse) to 10 (fully solved), how would you say the [targeted issue] problem has changed from 7 months ago to today?" To calibrate scores across regions, I removed any mention of media reports from the qualitative reports and gave the reports and accompanying evidence to an independent reviewer to re-score. The correlation between the original auditor score and independent reviewer score is 0.68, and Cronbach's Alpha for the two measures is 0.81. The outcome presented in the main text is a simple average between the original auditor and independent reviewer's scores, although the primary results are almost identical regardless of grader. Figure 5 shows the distribution of scores in the control group sample. The modal outcome is 0 (no improvement), the mean outcome is 2.3 (marginal improvement or planned improvement), and the standard deviation in the control sample is 2.7.

Independent audits and report card scores offer an attractive outcome measure for several reasons. First, the procedure was "double-blind" in the sense that auditors were not told about treatment assignment or purpose of the intervention, and community leaders and citizens were not aware that the independent audits were connected to the *WNS* investigations.[11] Another attractive feature of the design is that it focuses on concrete outcomes rather than citizen or leader perceptions. This distinction is especially important in the context of a media study in which reports themselves may have influenced citizen and leader expectations and satisfaction even in the absence of objective improvements in service delivery. Finally, the audit measure is easily to replicate or augment. Because the basis for evaluations is including in qualitative report cards, future analysts could easily create a separate metric for evaluating service delivery success and give a community an updated score. For example, the scores provided here emphasize a holistic assessment of service delivery improvements (giving points for repairs, partial improvements, and pledges). If researchers were interested in giving credit only for new and completed service delivery projects, the raw data is available for creating such a score for each community.

---

[11] Even if auditors had found out that a given community received *WNS* investigations and allowed the knowledge to influence their scores, any mention of local press coverage was removed from the audits before being passed to the independent reviewer.

Figure 5: **Distribution of Independent Audit Scores**, Control Sample
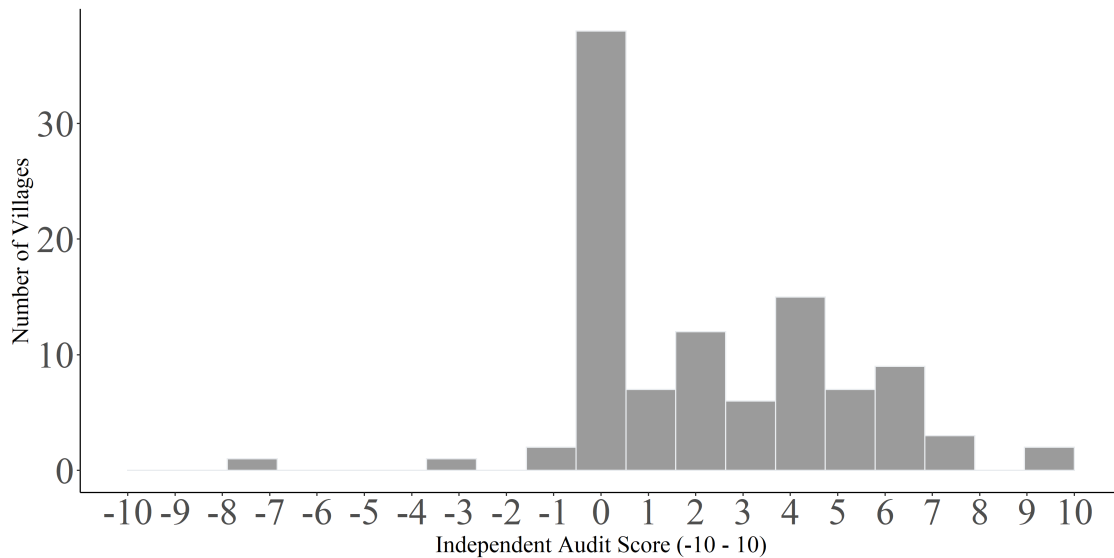
What do these scores mean in practice? To make the meaning of the report card scores more concrete, I can connect each auditor's quantitative score to their qualitative narratives about the community's experience. In communities that received a score of 0, service delivery improvements were predictably nil, or minimal progress was offset by comparable setbacks. In the village of Liwiti in the southeastern coastal region of Lindi, for example, citizens have been requesting assistance to repair a road ever since a memorable visit from the District Councilor, whose car got stuck in the mud and was forced to use a motorcycle to reach town, and the Tanzanian Urban and Rural Roads Agency (TARUA) in 2018. Over the 7 month project period, no progress was made towards repairing the road. In the rural village of Ngoji, the absence of piped water forces villagers to regularly dig shallow wells by hand or retrieve pond water that often makes them sick. Villagers dug two wells during the project period, but the wells only replaced wells that dried up in the same period. Stories like Liwiti and Ngoji are by far the most frequent outcome in the control group sample; 28% of control communities received a score of 0. Citizens told auditors that they wait for election season to receive even the promise of future support.

In communities that received scores between 1 and 3 citizens experienced positive but marginal improvements in service delivery, or plans for significant improvements were set in motion but not yet carried out. In the village of Ipagusa in Tanzania's central Morogoro Region, for example, more than half of the community lacks access to potable water because the local well is small and often broken. During the project period, the district government responded by investing just 60,000 TZS (approximately 30 US dollars) to clean dirt and grime around local wells, and did not drill additional wells to confront the severe water shortage. In Nyregete village in the Mbeya highlands, the local health clinic routinely runs out of supplies, the doctor is often unavailable because there are no staff quarters connected to the clinic and he lives in another village, and citizens from distant sub-villages need to walk 2-3 km

to access health services. The government announced plans to build a new clinic in 2023, but the only action taken during the project period was to construct a new staff quarters for the doctor. Ipagusa and Nyregete received a score of 1 and 3, respectively.

Communities that received audit scores about one standard deviation above the treatment mean (a score of 4, 5 or 6) experienced more tangible improvements in the targeted issue area. In Tutuo village of western Tabora district, which is home to over 1,500 students by only one school with ten classrooms and ten teachers, the government contributed 48 million TZS (about 20,000 USD) and citizens contributed labor to construct a new school. Construction of the school was underway but far from complete, and no action had been taken to resolve the teacher shortage. In Mabogini village at the base of Mt. Kilimanjaro, the absence of drainage ditches means that roads become flooded and impassible during the rainy season. During the project period, TARURA moved in to repair the roads: they laid gravel on the primary road and built two small roads to avoid the flooded area. Nonetheless, the new roads have severe potholes and still lack drainage ditches, so worries remain that the roads will flood next rainy season and soon fall into disrepair. Tutuo and Mkingala received a score of 4.5 and 5, respectively. In short, a one standard deviation improvement in a community's audit score relative to the control mean signifies the difference between the government promising development projects and the government actually initiating them, or the difference between a marginal response to a service delivery problem and a direct (albeit short-term) solution to it.

At the higher end of the audit score, a score from 7-10 means the difference between the government initiating a project and the government completing it, or the difference between a short-term solution to a service delivery problem and a more permanent solution. Just 5 communities in the control group (3%) received a score of 7 or higher. In Idunda village in the Udzungwa mountains, the scarcity of functioning water points forced villagers to rely on polluted water from the Ruli river. During the project period, the District government sent the Rural Water and Sanitation Agency (RUWASA) to repair broken water points, and the village council raised money to pay a local contractor to drill ten new water points in the Idunda and surrounding sub-villages. Idunda received a score of 7. One of the highest scoring control villages was Ismani in Iringa region, where the government responded to an under-resourced health center by contributing 400 million TZS (171,000 USD) to upgrade the health facility with a laboratory, a doctor's quarters, an operating room, and patient's ward. Villagers and villagers' relatives from outside the community also made significant donations. Idunda received a score of 9.

5.3.3 Responsiveness

After scoring overall improvements in service delivery, auditors assigned scores to the responses from specific actors. Auditors were also required to include specific statements about the activities of key actors in their narrative reports, and these narrative reports were re-scored by an independent reviewer.
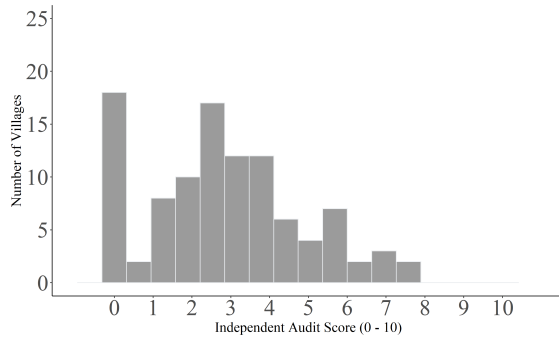
(a) Citizen Response Scores (0-10)
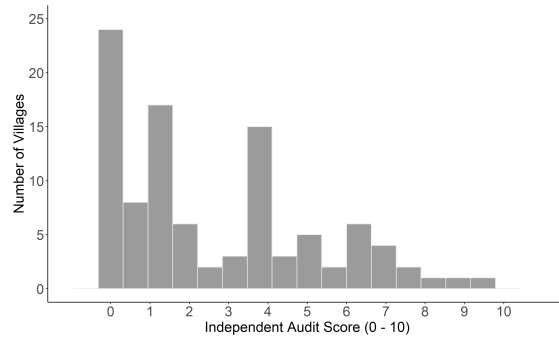
(b) Government Response Scores (0-10)

The first group auditors evaluated was citizens of the targeted community. Auditors were asked "on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by the citizens in this village/street to the X problem?" Auditors were trained to account for a broad range of citizen activities to resolve service delivery problems, including financial contributions, physical labor, and lobbying of government officials. The modal citizen responsiveness scores were 0 and 2.5 (17% of communities each), and the median score was 2.5. Citizens receiving scores between 1 and 3 held village meetings, lobbied government officials, or made limited financial or labor contributions, although they were often insufficient to resolve the problem entirely. For examples, villagers held community meetings about the problem issue in Tabora region (score = 1), filled potholes with sand as a short-term solution to broken road infrastructure in Geita region (score = 2), and lobbied ministry officials to request additional funds to resolve a water crisis in Tanga region (score = 2.5). Citizens earned scores between 4 and 7 if they made significant financial and labor contributions or enacted dramatic accountability measures, such as recalling an ineffective village leader. Citizens did not receive a score above a 7 in any community.

Auditors also scored the responses of the Tanzanian government. In a similar vein to citizen scores, auditors were asked "on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by the government to the X problem?" As Figure 6b makes clear, the government was more likely to take no or limited action than citizens: auditors scored government responsiveness a 1 or less in 42% of control communities. On the other hand, governments were also much more likely to significantly mobilize resources to fully address the service delivery problem, scoring 7 or higher in 10% of the control sample. After scoring the actions of the Tanzanian government as a whole, auditors scored the actions of specific sectors of government, including local elected officials, local bureaucrats, district-level ministry officials, and the Member of Parliament. As I discuss in more detail in the results section, government responsiveness to service delivery issues varies significantly from community to community and problem to problem.

5.4 Estimation

The estimand of primary interest is the intent-to-treat effect (ITT). Treatment "uptake" (communities where a journalist visited, produced, and broadcast a report) was 98.1%. As previously discussed, journalists were unable to complete reports in three communities at the request of government officials. These communities are still considered partially treated because reporters spoke with citizens and government officials, even though reports were never broadcast. Enumerators collected endline data in the three communities where *WNS* reports were forbidden.

The unit of observation for analyses are community leader respondents or direct observation measures taken at the community level. Intent-to-treat effects will be estimated using OLS regression. The regression model

$$y_i = \beta d_i + \gamma_1 pair_i + u_i$$

expresses the outcome y_i as a linear function of the randomly assigned treatment, binary variables for each of the community pairs (blocks), and an unobserved error term u_i. β represents the Intent-To-Treat (ITT) effect. The regression estimator is identical to the difference-in-means estimator, since the block variables are orthogonal to treatment assignment. I calculate p-values using randomization inference (RI) (2,000 permutations).

In keeping with my pre-analysis plan, the analysis of substantive outcomes also reports covariate-adjusted regression results. The LASSO procedure selects prognostic covariates from a set of administrative variables listed in Table 1. The number of selected covariates ranges from to zero to ten depending on the outcome, but due to the similarity across experimental groups at baseline, the estimates after adjustment closely resemble estimates without adjustment across all analyses.

6 Results

6.1 Compliance with Assigned Treatment

Was *Wakati Ni Sasa* implemented as planned? Table 2 shows the effect of treatment assignment on three first-stage outcomes, or manipulation checks: whether *WNS* reports were produced and broadcast, whether the reports were noticed and remembered by community government officials, and whether the reports influenced officials' beliefs about the likelihood of future media investigations.

Column 1 of Table 2 shows that *Wakati Ni Sasa* journalists produced and broadcast investigative reports in 100 of 103 communities (97%) assigned to the treatment condition.[12] In three communities, District-level government officials refused to cooperate with *WNS* journalists and forbade them from broadcasting their investigations. While *WNS* journalists were able to complete reports in some communities where District officials were first resistant by securing approval from higher-level Regional authorities, the research and implementation team determined that the

---

[12] An assigned research team member verified that each report was broadcast at least twice in the agreed-upon week.

risks to journalists were unacceptable in these three cases (see Appendix). Endline data collection was still collected in all three communities.

Did local leaders in targeted communities notice and remember *Wakati Ni Sasa* investigations? Enumerators asked three leaders in each community "how many times has a journalist visited this village/street in the past 7 months to address [targeted problem]" and "have you heard a radio report about [targeted problem] in your village/street in the last 7 months?" I coded the outcome as 1 if at least one official in the community remembered a journalist visit or heard a radio broadcast. While control communities were by no means saturated by local media reporting in the absence of *WNS*, local investigative reports were also not uncommon: at least one local leader recalled a journalist visit in 29% of control communities and reported hearing a radio report in 25% of control communities. Even so, *WNS* dramatically increased perceived press coverage in treated communities. On average, local officials in treated communities were 34 percentage points more likely to remember a journalist visiting their community and 28 percentage points more likely to remember hearing a news report about their community (one-tailed randomization inference p-value < 0.001). Overall, local officials recalled either a journalist visit or hearing a broadcast report in 75% of treated communities compared to 38% of control communities (one-tailed randomization inference p-value < 0.001).

I also observe evidence that *WNS* investigations influenced local leaders' beliefs about the likelihood of *future* media coverage. When enumerators asked village/street chairpersons "How likely do you think it is that a journalist will visit your village/street if it has a [targeted problem] problem in the future" 63% of leaders in the treatment group said a report was "very likely" compared to 46% in the control group. Taken together, Table 2 suggests that local media reports in the absence of *WNS* was common but idiosyncratic, while *WNS* investigations left a lasting impression on targeted local leaders. In other words, *WNS* reports increased the intensity of familiar styles of press-coverage rather than introducing a new accountability mechanism.

Table 2: **Effect of Treatment Assignment on Treatment Uptake**

	Reports	Local Official Awareness			Local Official Expectations
	Report Broadcast	Aware of WNS Index	Journalist Visit	Heard Report	Future Report Likely
	(1)	(2)	(3)	(4)	(5)
WNS Treatment	0.971	0.374	0.343	0.283	0.214
Standard Error	0.017	0.059	0.065	0.062	0.093
RI *p*-value	<0.001	<0.001	<0.001	<0.001	0.011
Hypothesis	+	+	+	+	+
Control Mean	0.00	0.38	0.29	0.25	1.31
Control Village SD	0.00	0.49	0.46	0.43	0.73
DV Range	0-1	0-1	0-1	0-1	0-2
Matched-pair controls	Yes	Yes	Yes	Yes	Yes
Lasso-selected controls	No	No	No	No	No
Adj-R^2	0.94	0.32	0.17	0.22	0.10
Observations	206	202	202	202	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. The dependent variable in Column 1 comes from monitoring data on whether a *Wakati Ni Sasa* report was completed and broadcast on the radio, and takes the value 1 if the report was broadcast at least twice and 0 otherwise. Column 2 reports an index of local official awareness of *Wakati Ni Sasa* reports which is the simple arithmetic mean of Columns 3 and 4. The dependent variable in Column 3 takes the value 1 if any community government official in the community answers "at least once" to the question "*How many times has a journalist visited this community or street to address the X problem in the past 7 months?*" The dependent variable in Column 4 takes the value 1 if any local official in the community answers "yes" to the question "*Have you heard a radio report about an X problem in your community in the past 7 months?*" The dependent variable in Column 5 is the community chairperson's response to the question "*How likely do you think it is that a journalist will visit your village/street if it has a [targeted problem] problem in the future?*" Responses took the value 2 if the chairperson responded "very likely" and ' if the chairperson responded "somewhat likely".

6.2   Impact on Public Goods and Services

I begin the analysis of results by considering the outcome of primary interest: change in the quality of public goods and services. In the *absence* of press reports, some targeted problems were addressed over the 7 month project period, but average improvements were small: the mean audit score in control communities was just 2.03 on a -10 to 10 point scale (standard deviation = 3.05), and auditors reported that the problem remained unchanged (37%) or had grown worse (8%) in nearly half of control communities.

How did *Wakati Ni Sasa* reports affect the quality of public goods and services? Table 3 and Figure 7 show that, communities targeted by *WNS* received audit scores that were 0.68 points higher than control communities on average, which amounts to more than one-fourth of a control group standard deviation. The one-tailed randomization inference *p*-value for the result is 0.03. Recall that a one standard deviation audit score improvement above the control group mean means the difference between a minimal response to the problem and a tangible short-term fix (e.g. gravel filling in an impassable road) or the initiation (but not completion) of a longer-term solution (e.g. breaking ground on a new clinic).
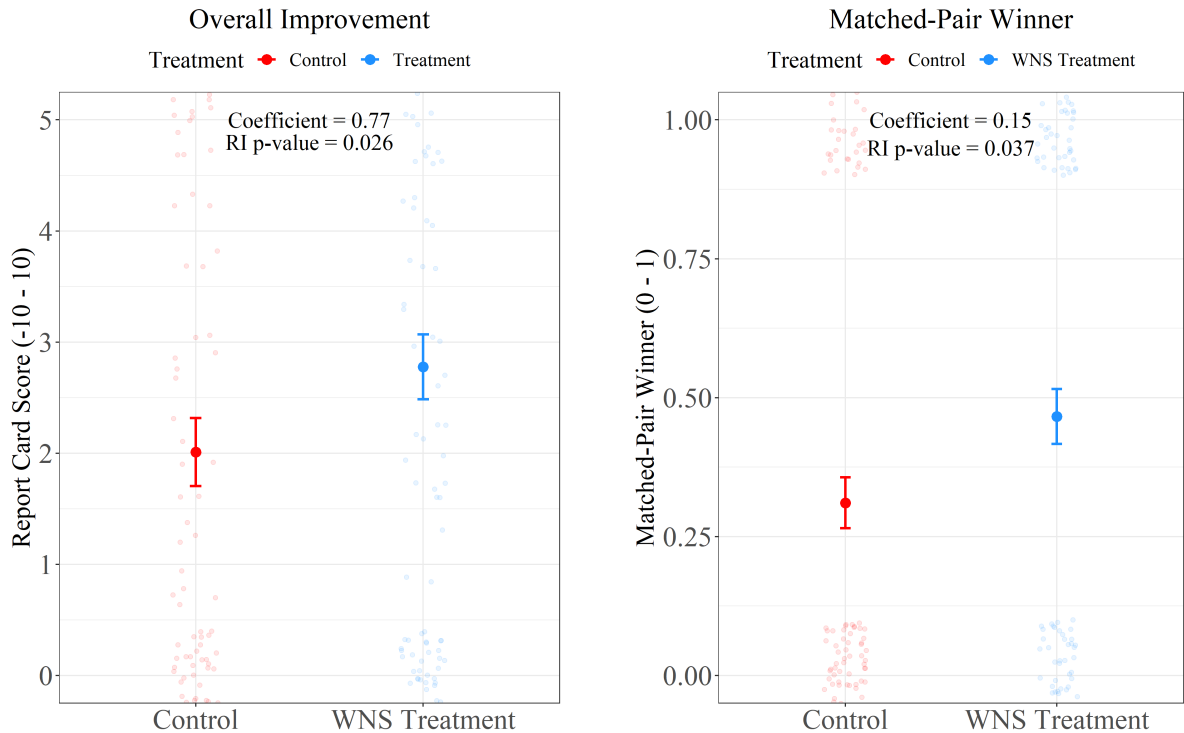
Figure 7: Treatment and control means for report card score and matched pair winner. Each light circle represents a village, dark circles represent treatment/control group averages. Matched pair winner has been recoded where 1 means village received higher score than matched pair partner, 0 means village did not receive higher score.

Another way to conceptualize the treatment effect is to ask how often treatment communities received higher audit scores than their matched-pair counterpart in the control group. Treatment communities scored higher than their control community counterparts in 50% of all matched pairs, tied in 13%, and scored lower in 35% (one-tailed randomization inference *p*-value = 0.03). The final pre-specified operationalization of the primary outcome is a simple binary indicator of whether the auditor recorded *any* improvement in the targeted issue area. Treated communities were about 7 percentage points more likely to see some improvement, although this result is only significant at the 0.1 level after adjusting for lasso-selected covariates.

Bear in mind that these estimates reflect the intent-to-treat effect of a community being targeted by a *WNS* report, and that the research team did not prevent radio stations from reporting on control communities for ethical reasons. Given that 38% of communities received press coverage of the targeted area during the project period, the effect of *WNS* reports relative to a pure control may have been considerably larger.

Table 3: **Primary Outcome** - Report Card Score

	Overall Score		Matched-Pair Winner		Any Improvement	
	(1)	(2)	(3)	(4)	(5)	(6)
WNS Treatment	0.684	0.684	0.155	0.155	0.068	0.091
Standard Error	0.361	0.361	0.086	0.086	0.067	0.063
RI *p*-value	0.033	0.033	0.028	0.028	0.112	0.077
Hypothesis	+	+	+	+	+	+
Control Mean	2.33	2.33	0.31	0.31	0.54	0.54
Control SD	2.77	2.77	0.47	0.47	0.50	0.50
DV Range	-10-10	-10-10	0-1	0-1	0-1	0-1
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	0	No	0	No	2
Adj-R^2	0.14	0.14	-0.59	-0.59	0.07	0.14
Observations	206	206	206	206	206	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Columns 1 and 2 come from an independent auditor's response to the question: "First, on a score of -10 (much worse) to 10 (fully solved), how would you say the [pre-specified problem name] problem has changed from 7 months ago to today?" Columns 3 and 4 report results from the same question, but where responses are coded 1 if the community had a higher report card score than its matched-pair partner, 0 if the community had the same report card score as its matched-pair partner, and -1 if the community had a lower score than its matched-pair partner. Columns 5 and 6 report results from the same question, but where responses are coded 1 if there was any improvement and 0 otherwise.

6.3 Robustness and Threats to Validity

The findings are robust to a variety of alternative outcome measures and model specifications. Table A1-Table A3 demonstrate that the treatment effects are similar whether I consider the score given by the original auditor or the score given by the follow-up auditor who only had access to evidence collected by the original auditor. The primary outcome measure is somewhat larger when using the first audit score alone. Appendix C also shows that the results are robust to a variety of alternative specifications, including the exclusion of matched-pair fixed effects and the inclusion of a battery of political, demographic, and economic covariates not chosen by the LASSO-selection procedure specified in the pre-analysis plan. Finally, Appendix C shows that the results are broadly consistent across topic areas, with the exception of health service delivery problems.

I also consider two threats to validity introduced by the possibility of interference between experimental units. First, *Wakati Ni Sasa* reports about treatment communities may have improved government responsiveness in near-bye control communities. For example, *WNS* reports may have catalyzed citizens in control communities who heard the reports to demand service improvements in their own areas. Similarly, government officials in control communities may have heard *WNS* reports, interpreted them as a sign that reports in control communities were more likely in the future, and improved their performance in anticipation of that eventuality.

Three observations help alleviate this concern. First, it is unlikely that the effect of *WNS* reports on the perceptions of citizens or government officials was dramatic. Table 2 shows that more than 30% of control communities were the target of local news reports even in the absence of *WNS*, and *WNS* reports were spread out between districts and over time so that no district would be subject noticeable surge in press attention. Second, I observe no evidence of spillover effects of *WNS* reports on government or citizen actions on other issue areas in treatment communities

(see Table A10), so it seems unlikely that the reports generated spillovers in different communities. Finally and most importantly, this form of interference would bias the estimated effect of *WNS* reports *downward* by improving public service scores in control communities.

The second threat to validity from interference between experimental units is government actions in treatment communities trading off with government actions in control communities. Again, it seems unlikely that such a trade-off would be dramatic, given that reports were spread about across districts and across time and that districts manage responses to as many as 40 communities. Moreover, district governments notoriously fail to spend down their allocated budgets, and it strains credulity to assume that ministries and district governments operate at full capacity (Carlitz 2017). Nonetheless, it is a weakness of the design that it cannot clearly definitively determine whether *WNS* reports increased the total amount of government effort and resources or only influenced the distribution of effort and resources. Considering the extreme case of a 1:1 trade-off between support for treated communities and support for control communities, the estimated treatment effect would be at most twice the size of the true effect. Estimating the precise nature of trade-offs between communities represents an important avenue for future research.

7   Mechanisms

How do local media investigations influence service delivery? Based on 12 months shadowing local journalists, I outlined six plausible channels of media influence on public service delivery (see Appendix): mobilizing citizens, attracting support from non-governmental organizations, helping citizens oversee local bureaucrats, helping citizens hold elected officials accountable, helping politicians hold bureaucrats accountable, and helping the ruling party monitor and sanction its own bureaucracy. However this "journalist-eye-view" of of media influence is subject to error, because journalists cannot observe community outcomes in the counterfactual world where they do not conduct an investigation.

Which channel of media influence predominates? In this section, I show that local news invstigations activate top-down channels of accountability *within* the bureaucracy. National and regional ministries are motivated to promote social welfare and avoid long-term threats to ruling party dominance (Magaloni and Kricheli 2010; Lorentzen 2014). Investigative news reports increase the likelihood that the grievances of local communities will reach Regional Commissioners (RCs) and regional ministry leaders, who are often described as the "eyes and ears" of the ruling party. Journalists help Regional Commissioners monitor and sanction ministry bureaucrats. These bureaucrats are keenly aware that public embarrassments may result in their demotion, firing, or transition to another region.

This section first shows that ministry officials change their behavior in response to local news reports, rather than citizens, NGOs, village-level bureaucrats, or elected officials. It then takes advantage of a negative shock to the sanctioning capacity of some Regional Commissioners - the shuffling of commissioners in 14 regions just as *Wakati Ni Sasa* was being launched - to show that media investigations are less effective when the likelihood of top-down

sanctioning is reduced.

7.1 Evidence from Report Card Scores

Having established a modest positive impact of *Wakati Ni Sasa* reports on overall service delivery outcomes, I now turn to the *Wakati Ni Sasa*'s effect on responses by specific actors. I first consider whether media investigations promoted responses by the Tanzanian government or citizens and civil society. Auditors were asked to score government, citizen, and non-governmental organization responses to the targeted problem on a scale of -10 (made the targeted problem much worse) to 10 (solved the targeted problem entirely).

In the control sample, citizen self-help (average audit score 2.9, score of 0 or less in 16% of communities) was more common than government support (average audit score 2.7, score of 0 or less in 25% of communities). In Mbangala village in the southeastern region of Lindi, for example, citizens contributed funding and labor to build the foundation for a new clinic. In Weruweru village at the base of Mount Kilimanjaro, villagers raised 32 million Tanzanian shillings (13,000 USD) to build the first four classrooms of a new school. In many control communities where government officials contributed funding and supplies, they did so only *after* citizens initiated service delivery improvements. Mbangala's Member of Parliament Mariam Kasembe, for example, contributed 100 bags of cement for additional bricks after villagers completed construction of the clinic's foundation.

Non-governmental organizations, on the other hand, were less likely to respond to service delivery problems than either citizens or the Tanzanian government. Nonetheless, non-governmental organizations provided crucial support to some control communities, especially around water issues. For example, the Islamic Foundation, a Saudi Arabian-funded community development organization based central Tanzania, sponsored the construction three water points in Dakawa village. Elsewhere in central Tanzania, the Israeli-based non-governmental organization Innovation:Africa provided technical support for construction of a piped water system connecting a primary village to two sub-villages.

How did *Wakati Ni Sasa* reports influence the relative contributions of government, citizens, and non-governmental organizations? Table 4 shows that *WNS* investigations reversed the control group's emphasis on citizen self-help relative to government support: *WNS* increased average government response by about 0.5 points (one-tailed RI p-value = 0.03) while *decreasing* average citizen responses by 0.36 points, although the latter result falls short of conventional statistical significance (two-tailed RI p-value = 0.180). *WNS* reports also did not appear to attract support from civil society: the effect *WNS* reports on non-governmental actions was small and statistically insignificant (coefficient = 0.11, one-tailed RI p-value = 0.290).

Table 4: **Government, Citizen, and NGO Response** - Report Card Score

	Government Score		Citizen Score		NGO Score	
	(1)	(2)	(3)	(4)	(5)	(6)
WNS Treatment	0.490	0.622	-0.325	-0.362	0.107	0.160
Standard Error	0.343	0.358	0.275	0.276	0.331	0.346
RI *p*-value	0.076	0.047	0.210	0.180	0.290	0.260
Hypothesis	+	+	two-tailed	two-tailed	+	+
Control Mean	2.54	2.54	2.83	2.83	1.63	1.63
Control SD	2.72	2.72	2.01	2.01	2.54	2.54
DV Range	-10-10	-10-10	-10-10	-10-10	-10-10	-10-10
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	7	No	2	No	7
Adj-R^2	0.18	0.22	0.10	0.14	0.18	0.15
Observations	206	206	206	206	206	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Columns 1 and 2 come from an independent auditor's response to the question: ""First, on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by the government to the X problem?" Columns 3 and 4 come from an independent auditor's response to the same question, except referring to citizens of the targeted community. Columns 5 and 6 come from an independent auditor's response to the same question, except referring to non-governmental organizations.

Turning to actors within the Tanzanian government, Table 5 shows report card scores for responses by community politicians, community bureaucrats, Members of Parliament, and ministries. There is no indication that *Wakati Ni Sasa* reports increased the average response of community politicians like village/street chairpersons (*mwinyikiti*) and ward councilors (*diwani*). This is not because community politicians lacked the resources or capacity to respond; local elected officials received the highest audit scores in the control group (average score 4.49). Their responses usually took the form of raising and distributing village funds, but also included rallying citizens to contribute labor to certain projects and lobbying higher levels of government for additional aid. However, local elected officials were almost entirely unaffected by *WNS*'s investigations on average (coefficient = 0.055, one-tailed RI *p*-value = 0.51).

Table 5: **Government Response Outcomes** - Report Card Score

	Local Elected		Local Bureaucracy		MP		Ministry	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WNS Treatment	-0.005	-0.005	-0.019	-0.019	-0.320	-0.320	0.573	0.573
Standard Error	0.254	0.254	0.249	0.249	0.312	0.312	0.346	0.346
RI *p*-value	0.506	0.506	0.500	0.500	0.828	0.828	0.048	0.048
Hypothesis	+	+	+	+	+	+	+	+
Control Mean	4.47	4.47	4.22	4.22	2.13	2.13	2.46	2.46
Control Village SD	1.97	1.97	1.89	1.89	2.60	2.60	2.81	2.81
DV Range	0-8	0-8	0-8	0-8	-5-9	-5-9	-1-10	-1-10
Matched-pair controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lasso-selected controls	No	0	No	0	No	0	No	0
Adj-R^2	0.02	0.02	0.10	0.10	0.12	0.12	0.26	0.26
Observations	206	206	206	206	206	206	206	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. These questions about specific government agencies are responses to the question: "Now we will ask you to give a score to different individuals and organizations in government. For each one, we are asking you to score only that person or organization's contribution to solving the X problem" Columns 1 and 2: "How do you score the Mwenyekiti's response?" and "How would you score the Diwani's response?"; Columns 3 and 4: "How would you score the Village Executive Officer or Ward Executive Officer response?"; Columns 5 and 6: "How would you score the response by the Member of Parliament?" ; Columns 7 and 8: "How would you score the response by the Ministry (examples: TARURA, TANESCO, RUWASA)?"

What about local bureaucrats? Village and Ward Executive Officers are charged with managing the imple-

mentation of local projects and overseeing the disbursement of village and district funds. According to the social accountability model of media influence, press coverage casts light on incompetence and malfeasance in local bureaucracies, helping citizens motivate service delivery providers to perform more effectively. Columns 3 and 4 offer sparse support for this story. Local bureaucrats are already among the most responsive actors in the absence of *WNS* reports (average score 4.25), but their activities were not discernibly affected by increased press attention (coefficient = -0.01, one-tailed RI *p*-value = 0.6).

WNS investigations also do not appear to influence Members of Parliament. There are some instances of Members of Parliament visiting sample communities and gifting bags of cement for clinic construction, new school supplies, or a diesel engine to power the local borehole. Nonetheless, according to report card scores MP's are are the least responsive actors in Tanzanian government: columns 5 and 6 show that MP's received the lowest average responsiveness score among all sectors of government (average score 2.12), and their efforts were slightly *lower* in communities that received *WNS* investigations (coefficient = -0.27, one-tailed RI *p*-value = 0.82).

Rather than influencing local elected officials, Members of Parliament, or local bureaucrats, it appears *Wakati Ni Sasa* investigations primarily triggered action by officials from the Tanzanian Rural and Urban Roads Agency (TARURA), the Ministry of Rural Water Supply (RUWASA), and the health and education ministries. In control communities, these ministry officials were nearly as unresponsive as MPs (average score 2.45). However, *WNS* reports increased average response scores by 0.67 points (RI *p*-value = 0.02), approximately one fourth of a control group standard deviation. Figure A1 shows that overall treatment effects were strongest across the three issue areas that fall most directly under the purview of ministry responsibilities: roads, water, and wildlife/environment.

7.2 Evidence from Heterogeneous Treatment Effects

Finally, I consider how variation in the oversight and sanctioning capacity of Regional Commissioners moderated the impact of *Wakati Ni Sasa* investigations. If *WNS* improved service delivery by activating top-down channels of accountability, we should expect *WNS* investigations to have been less influential in regions where systems of bureaucratic sanctioning were constrained.

To test this intuition, I exploit the decision by the new Tanzanian presidential administration, which took office just as the *Wakati Ni Sasa* program was launching, to demote or rotate Regional Commissioners in 14 Regions in May 2021 (the first *WNS* reports occurred in April 2021). The Administration's decision to rotate Regional Commissioners provides a useful proxy for decreased responsiveness by leaders of the regional bureaucracy. Regional Commissioners help the central government monitor local government performance (Katunzi and Spurk 2019), and, along with national ministers, are chiefly responsible for ensuring the smooth functioning of regional ministries. The newly selected Regional Commissioners did not start in their new positions until July 2021, when most of the *WNS* reports were finished, making it unlikely that they would have been positioned to hear and respond to *WNS*

investigations.

To test the possibility that *WNS* investigations activated bottom-up channels of electoral accountability, I also consider how the impact of *WNS* varied with opposition party strength.[13] If *WNS* was only influential in electorally competitive constituencies, it suggests that the influence of media investigations depend in part on bottom-up pressures familiar to many democracies. CCM controlled all wards and districts in the study sample in 2019. Nonetheless, I can assess underlying *opposition party strength* in each ward and district by drawing on election results from the 2014/2015 elections, in which opposition parties had a strong showing. In Table 6, "opposition party strength" is measured by the 2015 vote share of the leading opposition presidential candidate in a given constituency. The findings are nearly identical for outcomes from the parliamentary or presidential election and whether I use opposition vote share, CCM vote share, or margin of victory (see Appendix). Opposition presidential vote share ranges from from 0.52 (relatively strong opposition) to 0.15 (relatively weak opposition).

Table 6: **Heterogeneous Treatment Effects**, by New Regional Commissioner and Opposition Party Strength in 2015 Presidential Election

	(1)	(2)	(3)
WNS Treat	1.857	-1.439	0.438
	(0.723)	(1.686)	(2.199)
New RC	0.447		0.250
	(0.629)		(0.701)
WNS Treat * New RC	-1.651		-1.317
	(0.889)		(0.993)
RI *p*-value	0.065		0.187
Opp. Strength		-3.304	-2.787
		(3.161)	(3.473)
WNS Treat * Opp. Strength		5.942	2.213
		(4.458)	(4.905)
RI *p*-value		0.184	0.513
Blocked FE	No	No	No
Adj-R^2	0.021	0.008	0.035
Observations	206	204	204

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. The *p*-values for interaction terms are one-tailed tests in line with pre-analysis plan. Treatment is a binary indicator taking the value 1 if the village was assigned to the *Wakatni Ni Sasa* treatment. "New RC" takes the value 1 if President Samia Hassan assigend a new Regional Commissioner to the region in 2021. "Opp. Strength" is defined as the percentage of vote captured by opposition party Presidential candidates in the 2015 presidential election. Results are the same outcome measures reported in Table 3 and Table 5.

Table 6 reports heterogeneous treatment effects of *WNS* investigations on overall service delivery by (1) whether the community's Regional Commissioner was replaced in April 2021 and (2) opposition party strength in the community's district. Treatment effects on overall service delivery decrease substantially in regions where the RC was replaced, from 1.85 points on the -10 to 10 scale (*p*-value = 0.015) to 0.21 points (*p*-value 0.688). The randomization inference *p*-value of the interaction effect is 0.065. This provides suggestive, if not conclusive, evidence that the effect of *WNS* was conditional on the presence of a credible threat of sanctions from the Regional Commissioner. Treatment effects on overall service delivery also increase with opposition party strength (RI *p*-value = 0.184), although

---

[13]This test, along with tests for heterogeneous treatment effects by community remoteness and bureaucratic capacity, was included in the project's pre-analysis plan. The full results of these tests are included in the Appendix

the effect is not significant at conventional levels of statistical significance.

8   Discussion

An independent press is considered a bedrock of accountable governance in industrialized democracies, but its role in developing countries, especially those ruled by hegemonic party regimes, is less certain. This study examines the impact of local media investigations on the delivery of public goods and services in the "hard case" of Tanzania, where the reach of local media, the delivery of public services, and the extent of political competition is limited.

How can local media influence the provision of public goods and services under these circumstances? Drawing on more than 12 months of field work with radio stations and journalists, I developed a template for local media investigations modelled on the best practices of local reporters. We tested the effect of high-quality local media investigations through a national-scale field experiment conducted in partnership with 15 regional radio stations.

The primary finding is that local media investigations generated modest but meaningful improvements in the delivery of public goods and services 7 months after they were are aired. However, I observe no evidence that *WNS*'s influence was attributable to increased citizen self-help, NGO support, citizen mobilization of local service providers, or oversight of the bureaucracy by members of parliament. Instead, in keeping with theories of top-down intra-party accountability, only ministry officials appear to consistently respond to negative press reports. Qualitative evidence suggests that district level ministry officials are motivated to respond to the prospect of negative news coverage to prevent bad news from reaching their superiors and threatening their upward trajectory in the ruling regime. In line with theoretical expectations, *WNS* reports had almost no impact in regions without an operational Regional Commissioner to sanction lower level bureaucrats. Taken together, the results suggest a novel channel of media influence in hegemonic party regimes: helping ruling parties monitor and sanction their own bureaucratic apparatus.

These findings have important implications for scholarship on political accountability and distribution in hegemonic party regimes. First, they represent the first experimental evidence that investigative journalism improves government responsiveness outside of industrialized democracies. Evidence for the efficacy of information and social accountability campaigns outside fo developed democracies is decidedly mixed, especially with respect to tangible outcomes like service delivery (Grossman and Slough 2022). The results raise the intriguing possibility that media investigations, conducted by journalists with long-standing ties to the communities where they operate, activate channels of accountability that information and transparency campaigns do not.

Second, the findings highlight the importance of accountability relationships within ruling party bureaucracies. Extant scholarship on government responsiveness emphasizes the role of principal-agent relationships between citizens and politicians, politicians and bureaucrats, and citizens and bureaucrats. The findings suggest that in hegemonic party regimes, accountability relationships within the executive bureaucracy, and within the ruling party apparatus more broadly, may be especially important for understanding barriers to government responsiveness.

Finally, paper offers important lessons for policy. Funding local media investigations is extremely inexpensive: each *Wakati Ni Sasa* report cost less than 100 USD to conduct. Crucially, *WNS* reports engendered responses from government institutions rather than citizens or non-governmental organizations. As democratic institutions come under increasing strain around the world, identifying ways to support, strengthen, and sustain local news reporting offers a promising avenue for ensuring that governments remain responsive to the basic needs of their citizens.

References

Anderson, S. E., Buntaine, M. T., Liu, M., and Zhang, B. (2019). Non-Governmental Monitoring of Local Governments Increases Compliance with Central Mandates: A National-Scale Field Experiment in China. *American Journal of Political Science*, 63(3):626–643.

Arkedis, J., Creighton, J., Dixit, A., Fung, A., Kosack, S., Levy, D., and Tolmie, C. (2021). Can transparency and accountability programs improve health? Experimental evidence from Indonesia and Tanzania. *World Development*, 142:105369.

Banerjee, A., Enevoldsen, N., Pande, R., and Walton, M. (2020). Public Information is an Incentive for Politicians: Experimental Evidence from Delhi Elections. Technical Report w26925, National Bureau of Economic Research, Cambridge, MA.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., and Khemani, S. (2010). Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India. *American Economic Journal: Economic Policy*, 2(1):1–30.

Barbera, S. and Jackson, M. O. (2020). A Model of Protests, Revolution, and Information. *Quarterly Journal of Political Science*, 15(3):297–335.

Barro, R. J. (1973). The control of politicians: An economic model. *Public Choice*, 14(1):19–42.

Besley, T. and Burgess, R. (2001). Political agency, government responsiveness and the role of the media. *European Economic Review*, 45(4-6):629–640.

Besley, T. and Burgess, R. (2002). The Political Economy of Government Responsiveness: Theory and Evidence from India. *The Quarterly Journal of Economics*, 117(4):1415–1451.

Bhandari, A., Larreguy, H., and Marshall, J. (2021). Able and Mostly Willing: An Empirical Anatomy of Information's Effect on Voter-Driven Accountability in Senegal. *American Journal of Political Science*, n/a(n/a).

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*, 124(2):735–769.

Bobonis, G. J., Cámara Fuertes, L. R., and Schwabe, R. (2016). Monitoring Corruptible Politicians. *American Economic Review*, 106(8):2371–2405.

Bowles, J. (2020). Elite Preferences, Candidate Selection, and Local Public Goods: Evidence from Single-Party Tanzania. *Working Paper*, page 59.

Bowles, J. (2022). Identifying the Rich: Registration, Taxation, and Access to the State in Tanzania. *Working Paper*.

Brambor, T., Goenaga, A., Lindvall, J., and Teorell, J. (2020). The Lay of the Land: Information Capacity and the Modern State. *Comparative Political Studies*, 53(2):175–213.

Buntaine, M. T. and Daniels, B. (2020). Combining bottom-up monitoring and top-down accountability: A field experiment on managing corruption in Uganda. *Research & Politics*, 7(3):2053168020934350.

Buntaine, M. T., Hunnicutt, P., and Komakech, P. (2021a). The Challenges of Using Citizen Reporting to Improve Public Services: A Field Experiment on Solid Waste Services in Uganda. *Journal of Public Administration Research and Theory*, 31(1):108–127.

Buntaine, M. T., Zhang, B., and Hunnicutt, P. (2021b). Citizen monitoring of waterways decreases pollution in China by supporting government action and oversight. *Proceedings of the National Academy of Sciences*, 118(29):e2015175118.

Campante, F., Durante, R., and Tesei, A. (2022). Media and Social Capital. *Annual Review of Economics*, 14(1):null.

Carlitz, R. D. (2017). Money Flows, Water Trickles: Understanding Patterns of Decentralized Water Provision in Tanzania. *World Development*, 93:16–30.

Casey, K., Glennerster, R., and Miguel, E. (2012). Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*. *The Quarterly Journal of Economics*, 127(4):1755–1812.

Cheeseman, N., Matfess, H., and Amani, A. (2021). Tanzania: The Roots of Repression. *Journal of Democracy*, 32(2):77–89.

Chen, J., Pan, J., and Xu, Y. (2016). Sources of Authoritarian Responsiveness: A Field Experiment in China. *American Journal of Political Science*, 60(2):383–400.

Chwe, M. S.-Y. (2013). *Rational Ritual: Culture, Coordination, and Common Knowledge.* Princeton University Press.

Croke, K. (2021). The Impact of Health Programs on Political Opinion: Evidence from Malaria Control in Tanzania. *The Journal of Politics*, 83(1):340–353.

Dunning, T., Grossman, G., Humphreys, M., Hyde, S. D., McIntosh, C., and Nellis, G. (2019). *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I.* Cambridge University Press.

Durante, R., Pinotti, P., and Tesei, A. (2019). The Political Legacy of Entertainment TV. *American Economic Review*, 109(7):2497–2530.

Egorov, G., Guriev, S., and Sonin, K. (2009). Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data. *American Political Science Review*, 103(4):645–668.

Fearon, J. D. (1999). Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. /core/books/democracy-accountability-and-representation/electoral-accountability-and-the-control-of-politicians-selecting-good-types-versus-sanctioning-poor-performance/BA4FE99774A89B68A3E307F887A7BB49.

Ferraz, C. and Finan, F. (2011). Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. *American Economic Review*, 101(4):1274–1311.

Fox, J. A. (2015). Social Accountability: What Does the Evidence Really Say? *World Development*, 72:346–361.

Fukuyama, F. (2014). *Political Order and Political Decay: From the Industrial Revolution to the Globalisation of Democracy.* Profile.

Green, M. (2014). *The Development State: Aid, Culture & Civil Society in Tanzania.* Boydell & Brewer Ltd.

Grossman, G. and Michelitch, K. (2018). Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field Experiment in Uganda. *American Political Science Review*, 112(2):280–301.

Grossman, G., Michelitch, K., and Prato, C. (2020). The Effect of Sustained Transparency on Electoral Accountability. *Working Paper*.

Grossman, G., Platas, M. R., and Rodden, J. (2018). Crowdsourcing accountability: ICT for service delivery. *World Development*, 112:74–87.

Grossman, G. and Slough, T. (2022). Government Responsiveness in Developing Countries. *Annual Review of Political Science*, 25(1):annurev–polisci–051120–112501.

Guriev, S. and Treisman, D. (2019). Informational Autocrats. *Journal of Economic Perspectives*, 33(4):100–127.

Humphreys, M., Sánchez de la Sierra, R., and Van der Windt, P. (2019). Exporting democratic practices: Evidence from a village governance intervention in Eastern Congo. *Journal of Development Economics*, 140:279–301.

Humphreys, M. and Weinstein, J. M. (2012). Policing Politicians: Citizen Empowerment and Political Accountability in Uganda Preliminary Analysis. *IGC Report*, page 54.

Incerti, T. (2020). Corruption Information and Vote Share: A Meta-Analysis and Lessons for Experimental Design. *American Political Science Review*, pages 1–14.

Katunzi, A. and Spurk, C. (2019). Yearbook on Media Quality in Tanzania 2018. Technical report, Spurk Media Consulting Ltd.

Kosack, S. and Fung, A. (2014). Does Transparency Improve Governance? *Annual Review of Political Science*, 17(1):65–87.

Kruks-Wisner, G. (2022). Social Brokerage: Accountability and the Social Life of Information. *Comparative Political Studies*, page 00104140221089644.

Larreguy, H., Marshall, J., and Snyder, J. M. (2020). Publicizing malfeasance: When the local media structure facilitates electoral accountability in Mexico*. *The Economic Journal*, page ueaa046.

Lee, M. M. and Zhang, N. (2017). Legibility and the Informational Foundations of State Capacity. *The Journal of Politics*, 79(1):118–132.

Levitsky, S. R. and Way, L. A. (2012). Beyond Patronage: Violent Struggle, Ruling Party Cohesion, and Authoritarian Durability. *Perspectives on Politics*, 10(4):869–889.

Lieberman, E. and Zhou, Y.-Y. (2022/ed). Self-Efficacy and Citizen Engagement in Development: Experimental Evidence from Tanzania. *Journal of Experimental Political Science*, 9(1):46–63.

Lieberman, E. S., Posner, D. N., and Tsai, L. L. (2014). Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya. *World Development*, 60:69–83.

Lorentzen, P. (2014). China's Strategic Censorship. *American Journal of Political Science*, 58(2):402–414.

Magaloni, B. and Kricheli, R. (2010). Political Order and One-Party Rule. *Annual Review of Political Science*, 13(1):123–143.

Malesky, E., Schuler, P., and Tran, A. (2012). The Adverse Effects of Sunshine: A Field Experiment on Legislative Transparency in an Authoritarian Assembly. *American Political Science Review*, 106(4):25.

McCubbins, M. D. and Schwartz, T. (1984). Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science*, 28(1):165.

Mclellan, R. (2021). The Politics of Local Control in Electoral Autocracies. *Working Paper*.

Mdee, A. and Mushi, A. (2021). Untangling blame and responsibility for service delivery and local governance performance: Testing a grounded social accountability approach in Tanzania. *Local Government Studies*, 47(6):993–1013.

Morse, Y. L. (2012). The Era of Electoral Authoritarianism. *World Politics*, 64(1):161–198.

Morse, Y. L. (2019). *How Autocrats Compete: Parties, Patrons, and Unfair Elections in Africa*. Cambridge University Press, Cambridge ; New York, NY.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, 115(21):50.

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development*, 24(6):1073–1087.

Pan, J. and Chen, K. (2018). Concealing Corruption: How Chinese Officials Distort Upward Reporting of Online Grievances. *American Political Science Review*, 112(3):602–620.

Pettegree, A. (2014). *The Invention of News: How the World Came to Know About Itself*. Yale University Press, New Haven ; London, England, f first edition edition.

Przeworski, A., Stokes, S. C., and Manin, B., editors (1999). *Democracy, Accountability, and Representation*. Cambridge Studies in the Theory of Democracy. Cambridge University Press, Cambridge.

Pye, L. W. (1986). On Chinese Pragmatism in the 1980s. *The China Quarterly*, (106):207–234.

Raffler, P. (2019). Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from an Electoral Autocracy. *Working Paper*, page 77.

Raffler, P., Posner, D. N., and Parkerson, D. (2020). The Weakness of Bottom-Up Accountability: Experimental Evidence from the Ugandan Health Sector. *Working Paper*, page 126.

Reinikka, R. and Svensson, J. (2005). Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda. *Journal of the European Economic Association*, 3(2-3):259–267.

Sen, A. (2001). *Development as Freedom*. OUP Oxford, uk ed. edition edition.

Slough, T. (2020). Bureaucrats Driving Inequality in Access: Experimental Evidence from Colombia. *Unpublished Manuscript*, page 55.

Snyder, J. M. and Strömberg, D. (2010). Press Coverage and Political Accountability. *Journal of Political Economy*, 118(2):355–408.

Spurk, C. and Dingerkus, F. (2017). The need and the opportunities for sustainability – The case of local radio stations in Tanzania. *Journal of Media Business Studies*, 14(1):38–59.

Sturmer, M. (1998). *The Media History of Tanzania*. Ndanda Mission Press.

Truex, R. (2016). *Making Autocracy Work: Representation and Responsiveness in Modern China*. Cambridge University Press.

Tsai, L. L. (2007). Solidary Groups, Informal Accountability, and Local Public Goods Provision in Rural China. *The American Political Science Review*, 101(2):355–372.

Weghorst, K. (2022). *Activist Origins of Political Ambition: Opposition Candidacy in Africa's Electoral Authoritarian Regimes*. Cambridge University Press, Cambridge, United Kingdom ; New York, NY, new edition edition.

Appendix

A Six Theories of Media Influence

How might press coverage influence the quality of essential public goods and services like water, education, health care, and transportation? To identify channels of media influence, I spent more than 12 months in Tanzania prior to the launch of *Wakati Ni Sasa* working with local radio stations and shadowing investigative journalists. I also conducted in-depth interviews with more than 50 journalists, station managers, local politicians, and ministry officials across Tanzania. This preliminary qualitative field work helped identify six potential channels of media influence government responsiveness, which I then pre-registered as part of the *Wakati Ni Sasa* pre-analysis plan.[14] I discuss each in turn below.

A.1 Citizen Self-Help

The most direct channel of media influence on public goods provision is that coverage motivates citizens to engage in collective self-help. This channel of resonates with many "old-guard" media members in Tanzania because it reflects Julius Nyerere's philosophy that mass media's primary role is to educate and motivate citizens towards their own economic development (Mwaffisi, interview). Journalists shared three strategies for promoting citizen self-help: informing citizens about underappreciated problems in their area, persuading citizens that certain problems merit prioritization, and directly coordinating or advising appropriate citizen responses. A fourth strategy, generating "common knowledge" about a priority that is held privately by many citizens, is common in theoretical literature but was not emphasized by journalists (Barbera and Jackson 2020; Chwe 2013).

During a village visit to a peri-urban village suffering from water shortages in Morogoro Region, for example, independent journalist Mohammed Rajab spoke as much like a community mobilizer as a reporter, interspersing standard journalistic questions with exhortations to invest in private water pipelines. "We don't know when the government will come, my friends," Rajab reminded the crowd of 15-20 citizens who had gathered around the household where he was interviewing a key informant, "You need to work together!" The news item he later submitted for broadcast on Planet FM in Morogoro included reminders that government support for new water pipes would only be possible once citizens started digging trenches on their own.

Rajab's focus on local collective self-help is consonant with Elinor Ostrom's classic work on the co-production of public goods and services in settings with weak institutions (Ostrom 1996). While cooperative self-help is susceptible to classic coordination and collective action problems, Ostrom and others have documented a wide range of strategies that communities and local employ to overcome these challenges. Might local media represent one such strategy? Evaluations of *external* interventions to promote local self-help initiatives, often termed "community-

---

[14]See pre-analysis plan.

driven development", have yielded unimpressive results (Casey et al. 2012; Humphreys et al. 2019). Observational data on the influence of *media* on collective action is also mixed (Campante et al. 2022).[15] However, these studies focus on the effects of a community gaining *access* to new forms of media, rather than the effects of a community *being the subject* of press coverage. Local media coverage of type practiced by Mohammed Rajab in Morogoro has yet to be rigorously tested.

A.2 Non-Government Organizations

A second channel of media influence emphasized by both journalists and station managers is that media reports attract support from non-governmental organizations and philanthropists. This is an especially popular channel among media professionals in larger cities like Dodoma, Mwanza, and Arusha. In Tanzania's capital city of Dodoma, for example, a team of two journalists listed, off the top of their heads, more than 10 water-rights NGOs that they turn to when Tanzania's Ministry of Rural Water and Sanitation (RUWASA) is unresponsive. To attract support from civil society, radio journalists seek out dramatic personal-interest stories and convert them into social media posts (primarily Facebook and Instagram) that are most likely to be observed by NGO officials in urban centers. They also share this content directly with village chairpersons, ward councilors, and Members of Parliament, who use them as leverage to advocate for NGO support.

A robust literature examines the influence international and non-governmental development organizations in low-income countries, including the potential that public goods provision by these actors may undermine government accountability and long-term state capacity (Green 2014). If local media attracts support from NGOs but not governments, classic arguments for media's role in promoting government accountability would be undermined. To my knowledge, however, few studies have evaluated the influence of media reports on when and where NGOs provide development assistance.

A.3 Social Accountability (Citizen Oversight of Local Bureaucrats)

Media may also promote public goods provision by helping citizens monitor and sanction ineffective local service providers - contractors, doctors, teachers, and local bureaucrats. Many journalists have a favorite story of helping citizens bring a malfeasant or incompetent local bureaucrat to account. In theory, this "short route" to accountability offers a direct, cost-effective way for citizens to influence the design and implementation of public services (Kosack and Fung 2014). There is substantial observational evidence that communities which engage in local monitoring receive more and better services from the state (Tsai 2007). There is also some promising evidence that external interventions to support citizen monitoring of service providers generated substantial improvements in service quality (Björkman and Svensson 2009). By and large, however, interventions to support social accountability have not produced anticipated benefits (Olken 2007; Lieberman et al. 2014; Lieberman and Zhou 22ed) and have been susceptible

---

[15]Existing studies emphasize the distinction between traditional news media, which tends to facilitate collective action, and entertainment media, which tends to reduce social capital and political engagement.

to capture by local elites (Banerjee et al. 2010), while rigorous attempts to replicate the most promising models of social accountability have not been successful (Raffler et al. 2020). In Tanzania, an exceptionally careful evaluation of a social accountability intervention found that it produced no discernible effect service delivery or citizen satisfaction (Arkedis et al. 2021). The most optimistic interpretation of the existing base of evidence is that social accountability depends on external support for social accountability works only when paired with clear, direct, and actionable mechanisms for sanctioning under-performing bureaucrats (Fox 2015).

Media offers a promising tool for overcoming some of primary obstacles to effective social accountability. Like many of the interventions just described, news investigations may resolve information asymmetries between citizens and service providers and help citizens coordinate campaigns focused on specific service delivery issues and responsible actors. Journalists may also stimulate effective social accountability by providing clear templates for political action and acting as brokers by citizens and service delivery providers Kruks-Wisner (2022). There is some promising non-experimental evidence suggesting that local media improves service delivery through a social accountability channel Reinikka and Svensson (2005). However, therea are no experimental evaluations of independent media's contributions to social accountability.

A.4   Electoral Accountability (Citizen Oversight of Politicians)

Even when citizens do not directly monitor and sanction service delivery providers, they can still contribute to government responsiveness through their decisions at the ballot box (the "long route" to accountability (Kosack and Fung 2014). A journalist from Mbeya region said, "reports on those [service delivery] issues, can make [members of parliament] be scared they will not be selected if they are not responding." At the same time, many journalists emphasized that the influence of media was limited outside of election season. In the same interview, the Mbeyan journalist later explained, "…but you know, the MP has five years so maybe they are just making plans when they do not accomplish the assignment." When community politicians and members of parliament do make commitments to support local development projects, many journalists derisively refer to their statements as "*ahadi hewa*" (promises of the air).

In classic theories of democratic accountability, elections help citizens both *select* effective representatives and *motivate* their representatives once in office. For these systems of electoral accountability to be effective, voters need accurate and credible signals of politician's performance (Fearon 1999; Barro 1973). A robust literature finds that voters update their voting preferences in response to new information about about candidate malfeasance (Ferraz and Finan 2011) and overall performance (Bhandari et al. 2021; Bowles 2020), although evidence is mixed (Dunning et al. 2019; Incerti 2020). A smaller body of evidence suggests that incumbent politicians increase their effort when they know that voters are better informed about their performance (Grossman et al. 2020; Grossman and Michelitch 2018; Bobonis et al. 2016).

A functioning media environment has an obvious role to play in promoting electoral accountability in democratic countries. Both observational and experimental findings suggest that the threat of media reports motivate improved government responsiveness. However, media's favorable influence becomes less clear outside of industrialized democracies, where limited media reach (Larreguy et al. 2020), and more fragile electoral systems may undermine media's contribution to electoral accountability. Studies of transparency campaigns on politician behavior outside of industrialized democracies have met limited success (Malesky et al. 2012; Humphreys and Weinstein 2012).

A.5 Horizontal Accountability (Political Oversight of Bureaucracy)

A fifth potential channel of media influence on government responsiveness is helping elected officials to monitor and sanction service delivery providers. Politicians can improve the performance of service delivery providers by monitoring and sanctioning poor performance or by selecting better bureaucrats (McCubbins and Schwartz 1984; Slough 2020; Raffler 2019). Both strategies require politicians to be informed about the performance of service delivery providers, which depends on politicians having access to information about service delivery outputs and/or citizen preferences and grievances. In every investigation that I observed, journalists reached out to the relevant Diwani (Ward Councilor) and Mbunge (Member of Parliament) to share their findings and encourage the official to take action.

A growing body of empirical evidence evaluates programs that attempt to improve politicians' oversight of the bureaucracy. Olken (2007) finds that top-down monitoring of bureaucrats reduces corruption and increases performance of public works projects in Indonesia. Raffler (2019) finds that a program to train politicians to monitor bureaucrats in Uganda increased oversight behaviors and some measures of service delivery, although only in areas where politicians were not aligned with the ruling party. Slough (2020) argues that formal complaint systems in Colombia allow citizens to instigate political oversight of front-line service delivery providers.

However, there are significant barriers to improving political oversight of bureaucracies. Grossman et al. (2018) and Buntaine et al. (2021a) observe significant obstacles to citizens communicating their grievances to elected officials in Uganda. (Buntaine et al. 2021b) find that an initiative to help citizens report their problems to politicians in Uganda failed to improve waste service delivery because the quality of the information was low quality and did not fit into existing decision-making processes. Buntaine and Daniels (2020) find that a similar initiative in Uganda to failed to improve government responsiveness because politicians deliberately avoided knowledge of the monitoring in order to evade accountability.

A.6 Within-Party Accountability

Finally, journalists report influencing government responsiveness by activating accountability relationships *within* the ruling party hierarchy. Tanzania's executive bureaucracy is sprawling, and includes ministry officials

as well as regional, district, ward, and village-level executive officers. It is also deeply tied to, and often indistinguishable from, CCM party structures. Its internal promotion system is reminiscent of the cadre system in China (Chen et al. 2016; Pye 1986; Fukuyama 2014)

A distinguishing mark of a seasoned journalist in Tanzania is that they understand dynamics within the ruling party's local power structure. For example, the first thing Rodrik Makundi, a journalist in Kilimanjaro Region, does when he begins reporting on a story is to snap a picture or video and send it via WhatsApp to the relevant district executive officer and ministry official; this, he says, is even more crucial to his reporting process than the ultimate broadcast of the story. Makundi, like many Tanzanian journalists, talks about his job as a partnership with government officials. Sometimes, Makundi can get officials at TARUA (roads), TANESCO (electricity), or RUWASA (water) to respond to his stories with a video clip of the problem before the story is even broadcast. When these officials are unresponsive, he reaches out to Regional Commissioners and even Ministers. Makundi often plans his investigations to coincide with visits by major party officials to the region in the hopes of attracting their attention.

Rodrik Makundi's approach reflects a channel of media influence that is common in journalist's accounts of their own reporting process but sits uneasily with the academic literature on media and accountability in democracies. Rather than relying on the threat of electoral sanction, Makundi and other journalists form alliances with government officials, often sharing information about local governance failures and advice about how the government should respond. To elected officials, journalists provide a valuable mouthpiece for conveying their community's priorities to higher-level officials in the bureaucracy. To unelected ministry officials, journalists offer a valuable way to keep tabs on emerging challenges in their constituency. Ministry officials have an interest in CCM's continued success, but they do not hold elected office; their primary concern is signaling strong performance to regional and national superiors to continue their advancement within the ruling party. In a country where re-assignments of district and regional commissioners with regularity, they are especially keen to avoid embarrassment (Fatah, Interview).

The qualitative account of media's influence on intra-party accountability most closely resembles theories of media influence in autocratic regimes like China (Lorentzen 2014; Egorov et al. 2009). In these accounts, media facilitates top-down accountability by providing credible information to ruling elites about citizen grievances and the performance of local officials. Information from independent media may be especially useful when the state lacks the capacity to gather information itself or local officials misrepresent citizen grievances to their superiors (Chen et al. 2016). A number of recent studies suggest that civil society improves autocratic government responsiveness in China by providing bureaucrats with accurate information (Anderson et al. 2019; Buntaine et al. 2021b). However, the degree to which these findings hold in lower capacity governments like Tanzania remains very much in doubt.

B   Overview of Tanzanian Government
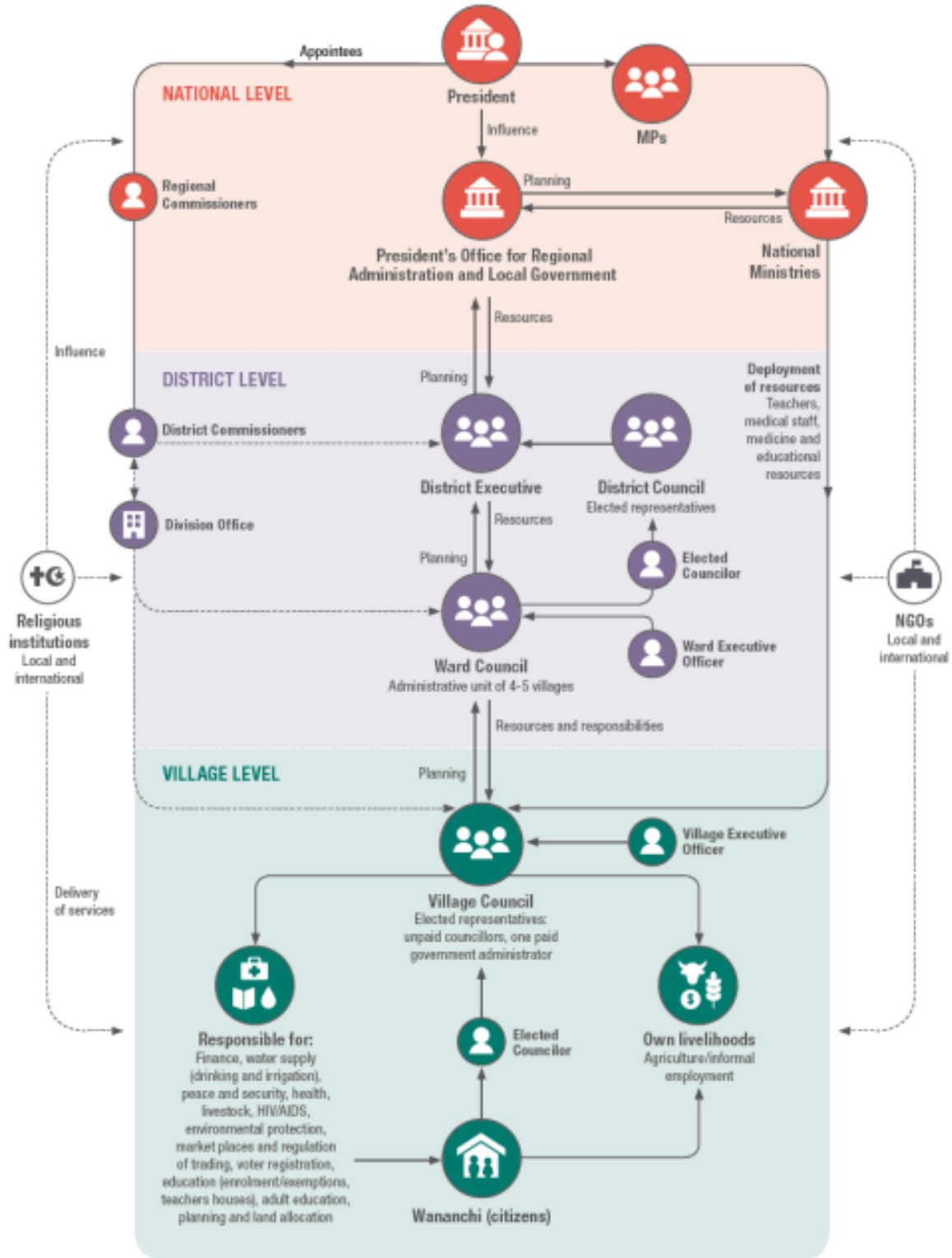
C   Supplemental Results

C.1   Comparison of In-Person and Follow-Up Audit Scores

Table A1: **Overall Improvement** - By Audit Scorer

	Combined Score		In-Person Audit Score		Blind Audit Score	
	(1)	(2)	(3)	(4)	(5)	(6)
WNS Treatment	0.684	0.684	0.311	0.311	0.068	0.091
Standard Error	0.361	0.361	0.172	0.172	0.067	0.063
RI *p*-value	0.031	0.031	0.028	0.028	0.112	0.077
Hypothesis	+	+	+	+	+	+
Control Mean	2.33	2.33	-0.16	-0.16	0.54	0.54
Control SD	2.77	2.77	0.87	0.87	0.50	0.50
DV Range	-10-10	-10-10	-10-10	-10-10	-10-10	-10-10
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	0	No	0	No	2
Adj-R^2	0.14	0.14	-0.95	-0.95	0.07	0.14
Observations	206	206	206	206	206	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Columns 1 and 2 come from a simple average of the next two outcomes. Columns 3 and 4 come from an independent auditor's response to the question: "First, on a score of -10 (much worse) to 10 (fully solved), how would you say the [pre-specified problem name] problem has changed from 7 months ago to today?" Columns 5 and 6 come from an independent auditor's response to thes same question based solely on the documentation submitted by the original auditor.

Table A2: **Government Response**, by Audit Scorer

	Combined Score		In-Person Audit Score		Blind Audit Score	
	(1)	(2)	(3)	(4)	(5)	(6)
WNS Treatment	0.490	0.622	0.427	0.564	0.553	0.608
Standard Error	0.343	0.358	0.390	0.401	0.346	0.344
RI *p*-value	0.076	0.047	0.143	0.087	0.050	0.039
Hypothesis	+	+	+	+	+	+
Control Mean	2.54	2.54	2.34	2.34	2.74	2.74
Control SD	2.72	2.72	2.99	2.99	2.71	2.71
DV Range	-10-10	-10-10	-10-10	-10-10	-10-10	-10-10
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	7	No	3	No	4
Adj-R^2	0.18	0.22	0.09	0.14	0.21	0.27
Observations	206	206	206	206	206	206

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Columns 1 and 2 come from a simple average of the next two outcomes. Columns 3 and 4 come from an independent auditor's response to the question: "First, on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by the government to the X problem?" Columns 5 and 6 come from an independent auditor's response to thes same question based solely on the documentation submitted by the original auditor.

Table A3: **Citizen Response**, by Audit Scorer

	Combined Score		In-Person Audit Score		Blind Audit Score	
	(1)	(2)	(3)	(4)	(5)	(6)
WNS Treatment	-0.325	-0.362	-0.350	-0.313	-0.301	-0.354
Standard Error	0.275	0.276	0.332	0.368	0.282	0.293
RI *p*-value	0.210	0.180	0.310	0.360	0.890	0.890
Hypothesis	+	+	+	+	+	+
Control Mean	2.83	2.83	3.17	3.17	2.49	2.49
Control SD	2.01	2.01	2.41	2.41	2.07	2.07
DV Range	-10-10	-10-10	-10-10	-10-10	-10-10	-10-10
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	2	No	8	No	5
Adj-R^2	0.10	0.14	0.12	0.16	0.08	0.09
Observations	206	206	206	206	206	206

 Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Columns 1 and 2 come from a simple average of the next two outcomes. Columns 3 and 4 come from an independent auditor's response to the question: "First, on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by citizens to the X problem?" Columns 5 and 6 come from an independent auditor's response to thes same question based solely on the documentation submitted by the original auditor.

C.2 Primary Results Without Matched-Pair Controls

Table A4: **Primary Outcome**, No Matched-Pair Controls

	Overall Score		Matched-Pair Winner		Any Improvement	
	(1)	(2)	(3)	(4)	(5)	(6)
WNS Treatment	0.684	0.684	0.155	0.155	0.068	0.091
Standard Error	0.361	0.361	0.086	0.086	0.067	0.063
RI *p*-value	0.033	0.033	0.028	0.028	0.112	0.077
Hypothesis	+	+	+	+	+	+
Control Mean	2.33	2.33	0.31	0.31	0.54	0.54
Control SD	2.77	2.77	0.47	0.47	0.50	0.50
DV Range	-10-10	-10-10	-10-10	-10-10	-10-10	-10-10
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	0	No	0	No	2
Adj-R^2	0.14	0.14	-0.59	-0.59	0.07	0.14
Observations	206	206	206	206	206	206

 Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns 1 and 2 come from an independent auditor's response to the question: "First, on a score of -10 (much worse) to 10 (fully solved), how would you say the [pre-specified problem name] problem has changed from 7 months ago to today?" Columns 3 and 4 come from an independent auditor's response to the question: "First, on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by the government to the X problem?" Columns 5 and 6 come from an independent auditor's response to the question "Next, on a score of -10 (made things much worse) to 10 (totally solved the problem), how would you score the overall response by the citizens in this village/street to the X problem?"

D Heterogeneous Treatment Effects

D.1 Electoral Competitiveness

Table A5: **Heterogeneous Treatment Effects**, by 2015 Opposition Party MP Vote Share

	(1) Overall Improvement	(2) Ministry	(3) Politicians
WNS Treat * Opp. Strength	3.158	-1.203	4.147
	(3.830)	(3.808)	(2.467)
RI *p*-value	(0.232)	(0.387)	(0.053)
WNS Treat	-0.615	0.919	-1.703
	(1.594)	(1.565)	(1.031)
Hypothesis	+	+	+
Opposition Strength	0.815	6.318	-8.103
	(17.174)	(12.701)	(4.795)
Blocked FE	No	No	No
Adj-R^2	0.05	0.23	0.13
Observations	204	204	204

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. THe *p*-values for interaction terms are one-tailed tests in line with pre-analysis plan. Opposition party strength is defined as the percentage of vote captured by opposition party MP candidates in the 2015 national election. Results are the same outcome measures reported in Table 3 and Table 5.

Table A6: **Heterogeneous Treatment Effects**, by 2015 Opposition Party Ward Vote Share

	(1) Overall Improvement	(2) Ministry	(3) Politicians
WNS Treat * Opp. Strength	-1.684	0.774	-0.463
	(3.773)	(4.247)	(2.079)
RI *p*-value	(0.334)	(0.411)	(0.439)
WNS Treat	1.466	0.157	0.030
	(1.566)	(1.784)	(0.819)
Hypothesis	+	+	+
Opposition Strength	-2.144	-2.742	-1.186
	(4.704)	(2.887)	(1.895)
Blocked FE	No	No	No
Adj-R^2	0.05	0.23	0.20
Observations	190	190	190

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. The *p*-values for interaction terms are one-tailed tests in line with pre-analysis plan. Opposition party strength is defined as the percentage of vote captured by opposition party Diwani (ward councilor) candidate in the 2014/2015 electoral cycle. Results are the same outcome measures reported in Table 3 and Table 5.

D.2 By Community Remoteness

Table A7: **Heterogeneous Treatment Effects**, by Distance from Secondary Road

	(1) Overall Improvement	(2) Ministry	(3) Politicians
WNS Treat * Distance	0.001	0.012	0.002
	(0.007)	(0.007)	(0.006)
RI *p*-value	(0.443)	(0.056)	(0.363)
WNS Treat	0.645	0.243	-0.101
	(0.460)	(0.433)	(0.300)
Distance	-0.002	0.019	-0.004
	(0.021)	(0.019)	(0.018)
Blocked FE	No	No	No
Adj-R^2	0.05	0.26	0.10
Observations	206	206	206

Note: * $p < 0.1$, ** $p < 0.05$, *** $p< 0.01$. Treatment is a binary indicator taking the value 1 if the village was assigned to the *Wakatni Ni Sasa* treatment. Distance from secondary road is measured in kilometers and calculated in R. Outcomes measures are identical to Table 6.

D.3 By LGA Efficacy

Table A8: **Heterogeneous Treatment Effects**, by LGA % Taxes Collected

	(1) Overall Improvement	(2) Ministry	(3) Politicians
WNS Treat * % Taxes	0.004	-0.010	-0.019
	(0.020)	(0.016)	(0.015)
RI *p*-value	(0.496)	(0.492)	(0.475)
WNS Treat	0.248	1.462	1.777
	(1.850)	(1.502)	(1.423)
% Taxes Collected	-0.222	-0.205	-0.046
	(0.057)	(0.013)	(0.061)
Blocked FE	No	No	No
Adj-R^2	0.06	0.23	0.13
Observations	204	204	204

Note: * $p < 0.1$, ** $p < 0.05$, *** $p< 0.01$. Treatment is a binary indicator taking the value 1 if the village was assigned to the *Wakatni Ni Sasa* treatment. LGA Efficacy is measured in by percentage of projected LGA tax collection actually collected in the according to the 2020-2021 Auditor's report. Outcomes measures are identical to Table 6.

Table A9: **Heterogeneous Treatment Effects**, by District Budget

|	(1) Overall Improvement	(2) Ministry	(3) Politicians
WNS Treat * log budget	0.437	0.358	0.809
	(0.560)	(0.494)	(0.366)
RI *p*-value	(0.160)	(0.247)	(0.027)
WNS Treat	-3.576	-2.917	-7.736
	(5.272)	(4.735)	(3.511)
Log per capita budget	6.487	6.222	1.272
	(1.608)	(0.459)	(1.751)
Blocked FE	No	No	No
Adj-R^2	0.07	0.23	0.15
Observations	204	204	204

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Treatment is a binary indicator taking the value 1 if the village was assigned to the *Wakatni Ni Sasa* treatment. District budget is the log of the total national budget allocation in 2020/2021 to the community's local government authority. Auditor's report. Outcomes measures are identical to Table 6.

E Conditional Average Treatment Effects

E.1 By Pre-Specified Problem Area

Figure A1 shows the standardized effect of the *WNS* treatment on overall report card score score and government response score, by pre-specified topic area. Treatment effects are concentrated among communities with transportation, water, and other (primarily environment and crime) service delivery problems, and are weaker for communities experienced education and health issues.

Figure A1: **Standardized Treatment Effect**, by problem type
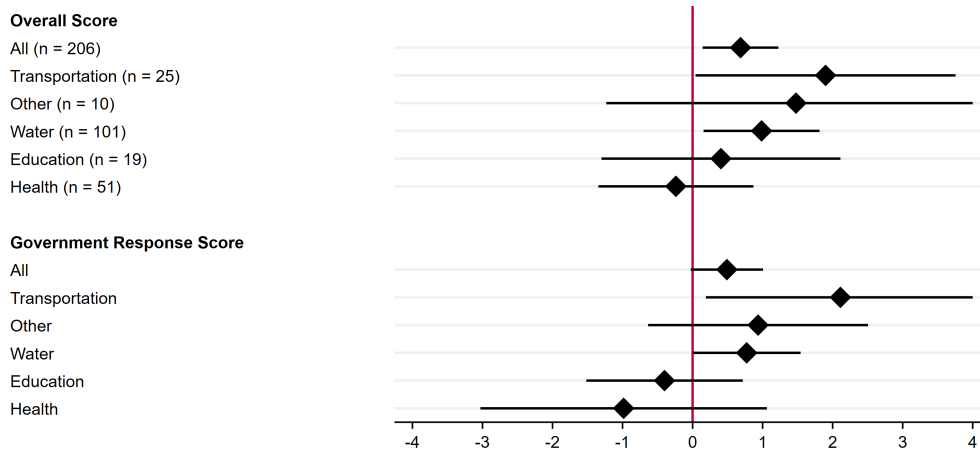
F Spillovers

To measure spillover effects, independent auditors also scored government and citizen actions on issue areas not targeted by *Wakati Ni Sasa Reports*. We find no evidence that *WNS* stimulated government or citizen actions on topics not covered by the investigative reports.

Table A10: **Government and Citizen Actions on Other Topics**

	Government		Citizen Actions	
	(1)	(2)	(3)	(4)
WNS Treatment	0.147	0.147	-0.146	-0.194
Standard Error	0.315	0.315	0.233	0.230
RI *p*-value	0.340	0.332	0.722	0.788
Hypothesis	+	+	+	+
Control Mean	4.27	4.27	0.75	0.75
Control SD	2.61	2.61	1.85	1.85
DV Range	-10-10	-10-10	-10-10	-10-10
Blocked FE	Yes	Yes	Yes	Yes
Controls	No	0	No	7
Adj-R^2	0.29	0.29	0.13	0.21
Observations	205	205	206	206

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns 1 and 2 come from an independent auditor's response to the question: "On a score of -10 to 10, how wuld you score GOVERNMENT's actions to address OTHER problems in their village (NOT XX problem) in the past 7 months?" Columns 5 and 6 come from an independent auditor's response to the same question but focused on citizen actions.

G   Do Citizens Care About Negative News?

I conducted a survey experiment to test a core assumption underlying the standard models of media influence. According to the standard account, media broadcasts are embarrassing to the ruling party because they spread news about governance failures beyond the borders of a particular affected community (Besley and Burgess 2001). To avoid these embarrassments, ministry officials and politicians might be especially motivated to respond to villages where they expect media reports to occur. But it is not obvious that citizens update their views of the national government and ruling party based on events they hear about in other communities. Citizens might not trust news stories from communities they are unfamiliar with, or they might place greater weight on their observations of the government's performance in their own community.

I implemented a survey experiment to test the influence of local news reports about government service delivery failures on citizens' opinions about community government, national government, and the ruling party. The primary hypothesis is that when citizens hear news reports about government service delivery failures in other communities in their region, they negatively update their views about the government and ruling party. The second hypothesis is that when negative news reports are followed by reports of a government response to remedy the service delivery failure, the negative impact on public opinion is attenuated.

The survey experiment is a two-armed intervention with a "positive" arm and a "no response" arm. Both treatment arms include a two minute news clip about a service delivery issue Tanga Region. The clip is a shortened version of a 15 minute news story that was broadcast on Tanga Kunani FM (TK FM) as part of the *Wakati Ni Sasa* radio program. The news item introduces water as an important issue in Tanzania, then introduces a journalist who is conducting investigations of water issues in Tanga Region. The journalist then shares a story about the lack of water service delivery in a near-bye village in Tanga and plays interviews with both villagers complaining about the impact that the lack of water has had on their health and safety, as well as an interview with a community government leader saying that the national government promised support but had not yet followed through.

After the initial news clip, the two treatment arms diverge:

- **Treatment 1 (Positive Response):** In the "positive response" report, the news follows up on the initial report by describing how the government successfully responded to villagers' needs and includes citizens and local officials thanking the government for their response.

- **Treatment 1 (No Response):** In the "no response" report, the report follows up the initial clip with the information that 5 months after the initial report, the government has still not responded to villagers' needs. The report indicates that the journalist will continue to investigate the effects of the water service delivery failure.[16]

---

[16]It is important to note that the survey experiment does not involve direct deception because respondents are never provided by inaccurate information. The "no response" report accurately reports that after 5 months, the government had not yet responded. The "positive response" report accurately states that after 6 months, the government resolved the water issue. After the outcome measures were collected, the surveyor played a final news clip about Tanzanian government water policy in Tanzania and actions in the village to ensure that all respondents received the full set of information by the end of the interview.

The two arms were designed to model the choice facing the Tanzanian government after the *Wakati Ni Sasa* report aired: they could respond to the problem, generating a "positive response" follow up report, or they could ignore the report, leading to a "no response" follow up report. My goal was to understand the public's reactions to these two outcomes.

Table A11: **Survey Experiment**, Positive Response Clip vs No Response Clip

	Beliefs	Attitudes	
	Posterior (1)	Nat. Gov (2)	Loc. Gov (3)
Gov. Response vs No Gov. Response	0.134	0.095	0.025
Standard Error	0.054	0.054	0.056
RI *p*-value	0.007	0.039	0.326
Hypothesis	+	+	+
Control Mean	0.29	0.59	0.49
Control SD	0.46	0.49	0.50
DV Range	[0-1]	[0-1]	[0-1]
Blocked FE	Yes	Yes	Yes
Controls	No	No	No
Adj-R^2	0.02	0.01	-0.00
Observations	309	321	324

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Posterior Beliefs [1 if Sometimes or Always, 0 Otherwise]: "Some people say that citizens receive the development support they need from government, and other people say citizens do not receive enough support. How about you, when citizens have development problems do you think the government always, usually, sometimes, rarely, or never responds?", Attitudes [1 if Very or Somewhat Satisfied]: "Now, I would like to ask you about how satisfied you are with different groups...[National Government, Community Government]"

Table A11 shows the effect of the "positive government response" news story relative to the "no government response" news story on respondent's beliefs and attitudes about the Tanzanian government. Respondents who listened to a story in which the government responded to a service delivery issue were 13 percentage points more likely to say that the Tanzanian government sometimes or always gives citizens the development support they need (RI *p*-value = 0.007) when compared to respondents who heard the "no response" clip. Beliefs about government responsiveness translated into more general attitudes towards the national government. Respondents who heard the "positive response" clip were 9 percentage points more likely to say they were satisfied with the national government (RI *p*-value = 0.039). Negative updating towards the government was focused more on national government than local officials; the treatment effect on attitudes towards community government was small and statistically insignificant.

Instead of comparing the two treatment arms (an investigation followed by no government response versus an investigation followed by a positive government response), Table A12 shows the impact of each treatment arm independently relative to a pure control.

Interestingly, the "no reponse" treatment does *not* influence citizens' beliefs about whether the government responds to the needs of citizens, but it *does* negatively influence attitudes about the national government. The "positive response" treatment had the inverse effect: it positively influenced citizens' beliefs about whether the government responds to the needs of citizens, but did *not* influence attitudes about the national government.

Table A12: **Survey Experiment**, Positive Government Response vs No Government Response

	Beliefs	Attitudes	
	Posterior (1)	Nat. Gov (2)	Loc. Gov (3)
No Gov Resp. vs Control	-0.013	-0.093	-0.065
Standard Error	0.082	0.081	0.083
RI *p*-value	0.401	0.037	0.116
Hypothesis	+	+	+
Yes Gov Resp. vs Control	0.120	0.002	-0.040
Standard Error	0.054	0.054	0.056
RI *p*-value	0.027	0.976	0.479
Hypothesis	+	+	+
Control Mean	0.31	0.69	0.56
Control SD	0.46	0.47	0.50
DV Range	0-1	0-1	0-1
Blocked FE	Yes	Yes	Yes
Controls	No	No	No
Adj-R^2	0.01	0.00	-0.00
Observations	463	484	493

Note: * p < 0.1, ** p < 0.05, *** p< 0.01. Posterior Beliefs [1 if Sometimes or Always, 0 Otherwise]: "Some people say that citizens receive the development support they need from government, and other people say citizens do not receive enough support. How about you, when citizens have development problems do you think the government always, usually, sometimes, rarely, or never responds?", Attitudes [1 if Very or Somewhat Satisfied]: "Now, I would like to ask you about how satisfied you are with different groups...[National Government, Community Government]"

H   Distributional Effects

Did *Wakati Ni Sasa* reports generate dramatic improvements in a few communities or modest improvements across many communities? Figure A4 shows a smoothed distribution of audit scores for treatment (blue) and control (red) communities. Communities assigned to receive *WNS* investigations were more somewhat more likely to receive positive scores ranging from 2 to 10, and were less likely to receive zero or negative scores. While it it is impossible to rule out the possibility that *WNS* reports inspired large-scale responses in some communities that were then offset by negative responses in others, Figure A4 lends credence to the interpretation that *WNS* had modest-to-moderate influence across many communities rather than large scale influence in a few. *WNS* may also have forestalled the *deterioration* of public goods and services (reflected in a negative audit score), which occurred in 8 control communities but just 2 treated communities.

Figure A2: **Smoothed Distribution of Independent Audit Scores**, Treatment versus Control Comparison
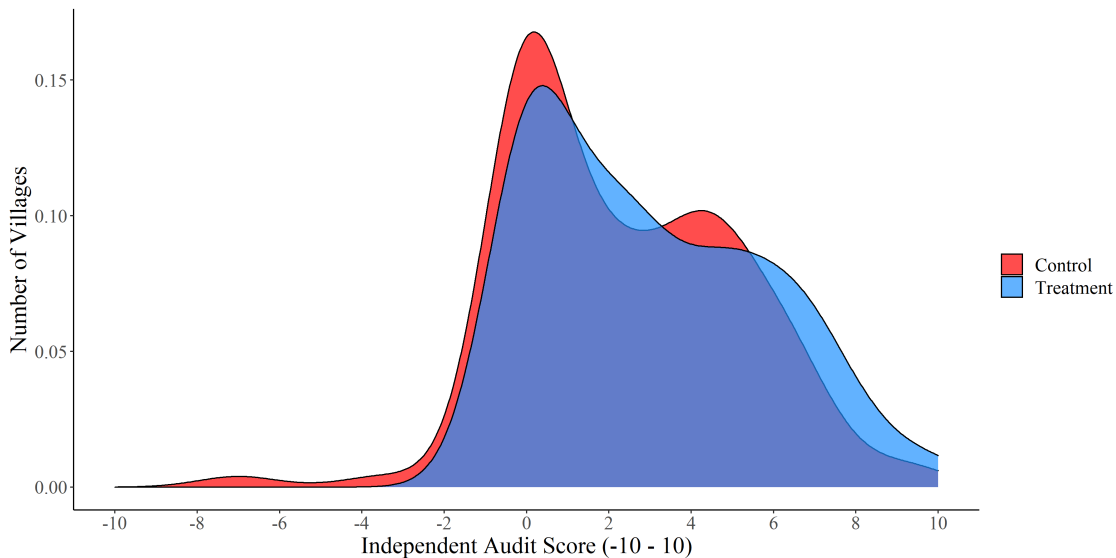

Figure A3 and Figure A4 offer two additional ways to visualize this finding. By-and-large, we do not observe substantial differences in the treatment effect across quantiles, suggesting that the treatment did not have significant distributional effects (eg effecting communities that tended towards less government responsiveness more than communities that tended towards higher government responsiveness).

Figure A3: **Cumulative Density of Audit Scores**, Treatment versus Control Comparison
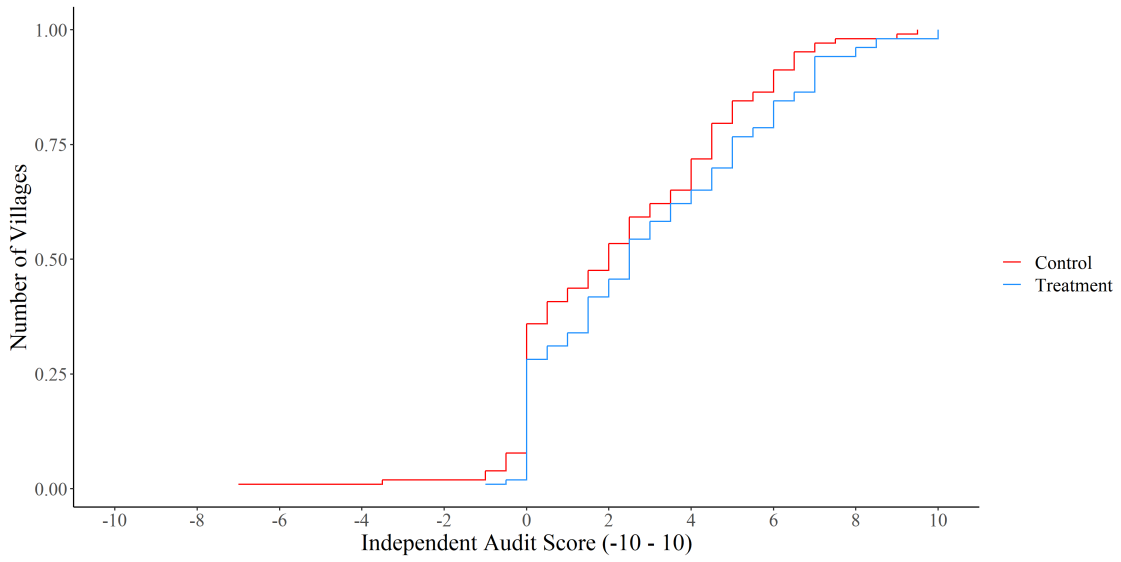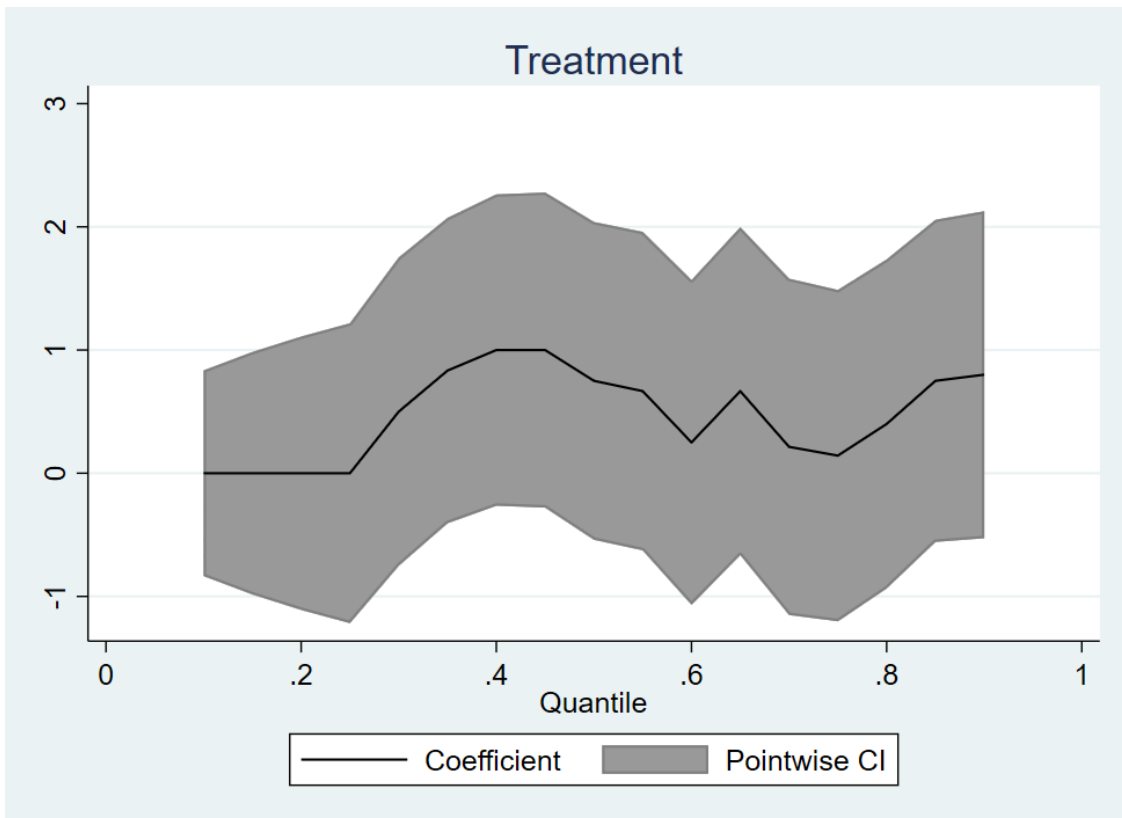

Figure A4: **Quantile Regression Plot**, Treatment versus Control Comparison

I Power Calculations

I performed power analyses for three measures of the primary outcome: the report card score for primary service delivery, the matched-pair winner of report card score, and a binary indicator. Table A13 shows minimal detectable effects (MDE) in standard deviation units using input parameters from the dataset for 80% power at α = 0.05 (one-sided test), as well as the estimated average treatment effect for reference. The explanatory power of matched-pari blocks is 0.55.

Table A13: Minimal Detectable Treatment Effects

Outcome	MDE (Power=0.8)	ATE
Overall Response (-10-10)	0.69	0.684
Overall Response (Matched Pair)	0.306	0.311
Overall Response (Any)	0.125	0.06

J Ethics

J.1 Equipoise

The principle of equipoise holds that an experimental intervention is ethically justifiable only if there is reasonable uncertainty among researchers that any arm of the intervention may be better for experimental subjects than any other arm. Recently, MacKay has introduced the term "policy equipoise", which holds that all treatment arms should also be in a state of equipoise with the best reasonably attainable alternative intervention.

The uncertainty around the effects of media, accountability, and community empowerment programs in low-income countries suggests that the current project satisfies the equipoise standard. There are no experimental studies of the effects of investigative media reports on government responsiveness, and more general studies on the effect of transparency programs have produced mixed results (?). Just as importantly, there is no consensus around better *alternative* interventions to address government service delivery failures at a similar cost and scale to this study's intervention (+/- 125 USD per village). Like media interventions, community empowerment programs and information treatments have produced mixed and/or modest results.

In short, current scholarly equipoise around the appropriate bundle of interventions to address government responsiveness is the central ethical justification for this project. A village receiving a media report is not obviously better off than a village that does not receive a media report, and neither village is obviously better off than a village that receives any other intervention attainable at a similar cost.

J.1.1 Scarcity

Even in the absence of equipoise, the study could be justified on the grounds of resource scarcity. Given a budget constraint by the radio station and research team, it would have been impossible to support media reports in all 200 villages. Even in the absence of policy equipoise, resource scarcity can justify randomization if no participant has a stronger moral claim to the intervention than other participants. We did consider the possibility that certain villages had a unique moral claim to media coverage relative to the rest of the sample based on the nature of their problem.

However, two considerations cut against this concern (apart from the equipoise argument from above). The first is that it would have been very difficult to differentiate the villages in "extreme need" from other villages in the sample in the absence of further investigation. All villages that radio stations identified expressed a strong need for government support. The second consideration is that the study design did not prohibit the village from receiving a media report treatment from the partner radio station or any other media outlet. If a village's situation proved particularly dire, there were still options for delivering the intervention, albeit outside the scope of this study.

Nonetheless, the research team was attentive to any village reports proposed by radio stations that seemed to demand urgent attention from the station, such that it would have been unethical to assign the village to the control group. In one case (a story of ongoing evictions and physical abuse of citizens by police), the research team

recommended removing the village from the sample and arranged support for a media report outside the research design. The removal of this village occurred prior to random assignment.

J.2 Researcher Role with Respect to Implementation

This study is not a case of a researcher independently evaluating an intervention that would have occurred in the researcher's absence: the investigative media reports would not have occurred when, where, and how they did in the absence of the researcher's participation and financial support. This level of researcher involvement in the intervention design implies a unique ethical responsibility by the researcher to the participants in the study.

The research team spent 1.5 years reaching out to radio stations and building relationships with station managers and reporters to understand the scope of their reporting, safety and ethical concerns, and their interest in a large-scale research project. The investigative report template was developed in collaboration with media houses and journalists, and radio stations held primary control and ultimate veto power over where to conduct reports, whom to interview, and what to include the broadcasts. Radio stations controlled the sampling process and carried out the reporting without researcher interference. The research team, in turn, provided funds for transportation, per diem, and production totally +/- 125 USD per report. These funds did not cover the full expenses of the reporting, which were shared by the media houses. The research team also reviewed all episodes before they were broadcast to ensure that no content would be aired that would put the radio station, journalists, communities, or research team at risk. Ultimately, no content was censored by the research team, although some recommendations were made to interview government officials for "balance" that is legally mandated by the Tanzanian government.

It is important to note that in the absence of the researcher's involvement, the radio stations would still have reported, produced, and broadcast local news stories. However, the researcher's involvement undeniably shaped the nature of the reports in three ways. First, transportation support allowed journalists to visit villages further away from the station than they would otherwise generally visit. Second, the production support enabled reporters to carry out more thorough investigations than standard news reports. Finally, the research design meant that radio stations produced more local news stories than they would typically produce.

J.3 Potential Harms to Participants or Non-participants

There are three categories of participants in this research project.

J.3.1 Journalists and Media Houses

The first category of participants is participating media houses and journalists. In the worst case scenario, journalists could have been harassed, arrested, or otherwise harmed by communities who did not want certain topics to be broadcast or aired. We took the following steps to minimize potential risks to media houses and journalists.

- The research team engaged in 1.5 years of planning conversations with media houses to ensure that the design of the intervention would not put them at legal risk.

- The research team agreed that all reports would include opportunities for government responses, in line with Tanzanian media law.

- The research team also agreed to exclude topics that media houses felt would invite government reprisal. These topics varied from region to region. In Mtwara Region, for example, we did not pursue stories related to sensitive issues around refugees entering the country from Mozambique.

- The research team made clear to journalists that they could abandon an investigation with no financial penalty if they felt that continuing to report would put them at risk, which they chose to do in two cases.

- The research team gave ultimate veto power over whether and when to air broadcasts to media houses.

- For all reports, the radio station invited government officials to respond on the record to citizen concerns in the report.

- The research team elected not to implement the project in the run-up to Tanzanian presidential election to avoid the perception of interfering with the electoral process.

J.3.2   Community Members in Treatment Areas

The second category of participants is community members in treatment areas. It is important to note that individual community members who elected to speak to journalists are not technically research subjects, in that no research data was collected from or about them. Nonetheless, it is easy to imagine a scenario in which the research project resulted in harm to community members who spoke with the media for the reports. For example, they could have experienced retribution from community members for speaking out about controversial issues. We took the following steps to minimize potential risks to community members who collaborated with journalists and media reports:

- No community member was forced to participate in an investigative report or was named in the report without their consent.

- Any community member was free to request that their contribution be considered off the record so that their name and identifying information was not associated with the reports

- All community members were given the contact information of participating journalists and allowed to follow up if had concerns of questions about the reports.

- The research team was in weekly contact with every media house for updates on any threats to participant safety. There were no reports of adverse consequences for participating community members.

It is also plausible that the reports may have led to government responses that left community members in participating villages worse off, whether or not they were interviewed by journalists. Uncertainty around the net effects of the reports on treatment communities is part of what generates equipoise, discussed in the sections above. If we had observed systematic negative consequences for participating communities, we would have closed the project. However, no such consequences were observed.

J.3.3 Government Officials in Treatment Areas

The third category of participants is government officials in treatment areas. The actions or inactions of government officials were often the subject of media reports. Reports may therefore have had negative consequences on the careers of elected officials and bureaucrats, although more severe harm was deemed unlikely and was not observed during the research period.

We ultimately deemed negatve career outcomes for government officials an acceptable negative risk, in line with interventions to promote accountability of bureaucrats and elected officials. In line with Tanzanian media law, government officials at the village, ward, and district level were in all cases invited to respond to media reports (on or off the record) and invited to follow-up with the media house after the report was broadcast.

J.3.4 Community Members in Control Areas

The final category of individuals potentially harmed by the project is community members in control areas. There are two plausible pathways by which individuals could be harmed. First, if a given region has limited financial resources, media reports in one village could have resulted in the diversion of resources from control villages towards treatment villages (also a violation of the non-interference assumption). Second, if media houses were focused on treatment areas, they may have ignored important social concerns in control areas to the detriment of communities in those areas.

The research team took several steps to avoid these concerns:

- The research team explicitly informed media houses that they were *not* restricted from reporting on issues in control areas. This avoided the risk that serious harms occuring in control areas would be ignored.

- The research team conducted the study *after* the budget process was underway. This meant that the primary focus of the intervention was on mobilizing action by at the Ward level (the unit of randomization) rather than the District or Regional level (where budget tradeoffs would be more likely to negatively impact control areas.

Nonetheless it is impossible to discount the possibility that media reports on one area may result in less attention on other areas. On the other hand, it is possible that media reports on one area may have increased attention on other areas as political actors sought to preempt the possibility of future reports. This uncertainty is part of what motivates and ethically justifies the research design.

J.4 Potential Harm from Data Collection or Research Protocols

J.4.1 Informed Consent for Randomization

Because of the nature of the intervention, community members were not given informed consent over the randomization process. As described in the section above, however, consent was implied by voluntary participation in the reports in much the same way that consent is implied by voluntary participation in community meetings during community development interventions

J.4.2 Informed Consent for Interview Participation

Data collection was conducted independently of the intervention and the enumeration team was blind to the treatment status of villages. All subjects will be provided with both written and oral explanations of the consent process and given the option to refuse participation. All data will be collected on password protected tablets using SurveyCTO/Open Data Kit software. Upon completion of the the survey, survey data is synched to an encrypted SurveyCTO server and downloaded to a laptop held by the research team and encrypted using Boxcryptor.

J.4.3 Community Consent

Survey enumerators will secure consent from Regional, District, Ward, and Village officials before conducting any surveys. This consent is both legally mandatory and ensures the safety of enumerators and research participants.

J.5 Potential Harms to Field Staff

Because of the geographic coverage of the study, field staff will be required to travel long distances, sometimes on public transportation or motorcycle. The research team takes the health and safety of staff extremely seriously. All staff are provided with medical insurance in case of accidents. All staff are provided per diem sufficient to cover overnight stays in safe lodgings so that they do not need to travel at night. All staff are paid salaries in compliance with Tanzanian employment law and their salaries are reported to government.

Additional risks introduced by the COVID-19 pandemic are discussed in an independent section below.

J.6 Financial or Reputational Conflicts of Interest

The research team reports no financial or reputation conflicts of interest in this study.

J.7 Intellectual Freedom

The research team reports no contractual limitations to reporting the results of the study. The Tanzanian government requires that any research be submitted to their research compliance office before publication, but there are no restrictions on what findings can be published.

J.8 Feedback to Participants or Communities

Because this study involves collaboration with radio stations, there is a in-built opportunity for sharing research findings with participating communities. Radio stations have agreed to broadcast reports about the findings, which will enable treatment, control, and non-sampled communities to hear about the results.

J.9 COVID-19

This project was implemented and data was collected in the midst of the COVID-19 pandemic. The research team takes its responsibilities related protection of research staff and subjects seriously. We obtained approval from Columbia University's COVID-19 review board to carry out the data collection, and designed transportation and data collection procedures with COVID-19 risks in mind. Shortly before the start of endline data collection, the Omicron

variant of COVID-19 was discovered, so endline data collection was converted to phone interviews.