# Pre-Analysis Plan: Do At-Scale Social Audits Improve Service Delivery of a Large Public Works Program?

December 3, 2021

**Abstract**

In this study, we intend to study the recent rapid expansion of Bihar's social audits initiative as a measure to improve access and delivery of social protection under India's national workfare scheme. We undertake a causal evaluation utilizing administrative data from a randomized controlled trial across the state. Initial analysis using the administrative data finds that audits lead to a significant, sustained reduction in work as reported in administrative data, a result that could reflect some combination of reduced corruption — a goal of audits, or declines in work provision — an unintended consequence. Looking ahead, we aim to collect additional survey data needed to differentiate between these possibilities and provide further direction for program improvement.

# 1 Introduction

Well-functioning social protection programs can play a key role in alleviating poverty, fostering resilience, and reducing inequality. These programs have an especially important role to play in low- and middle-income countries, yet many suffer from challenges including weak governance, mistargeting, leakage, and partial implementation of program features (Olken and Pande, 2012; Brown et al., 2018; Ravallion and Chen, 2008).

One theoretically appealing tool to address implementation problems is "bottom-up", community-led monitoring of program implementation, where potential beneficiaries have both strong incentives to advocate for improved program quality as well as the information required to assess implementation issues (Finan et al., 2017; Molina et al., 2016). An alternative is to take a "top-down" approach, in which higher-level agents of the state lead program oversight. While this approach is effective in some limited-capacity settings, other evidence indicates that there is scope for a top-down approach to backfire, especially when those doing the monitoring can be corrupted by those being monitored (Finan et al., 2017).

Against this backdrop, we study the impact of "social audits" of one of India's largest social protection programs, the Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) in the Indian state of Bihar. MGNREGS guarantees at least one hundred days of unskilled manual labor to every rural household that demands work; the program is self-targeted and aims to provide income support to the rural population. Social audits of MGNREGS combine community-led monitoring with external oversight: an external team of mostly female auditors reviews the program's implementation, and the audit culminates in a public meeting where local residents are supposed to discuss the results of the audits and hold local officials accountable. We aim to answer if this combined top-down external monitoring and bottom-up accountability system, improve implementation and accountability. To identify the causal impact of social audits on MGNREGS program performance, we collaborated with the state agency responsible for implementing social audits to conduct a randomized controlled trial (RCT) that leveraged the state's ambitious audit expansion plan in the 2019-2020 fiscal year.

Our initial analysis of the audit impacts focused on administrative data outcomes, in large part because the Covid-19 pandemic made it impossible to conduct in-person follow-up surveys. This document describes the strategy adopted for that analysis (which was conducted prior to publishing a pre-analysis plan) and outlines the additional data we expect to collect in a follow-up survey to be conducted in early 2022. Here we pre-specify how we plan to conduct additional analysis, with a focus on the survey data we aim to collect and hypotheses we will test.

The RCT results from the administrative data show that audits reduced the number of days worked under the program, and by extension reduced overall program expenditure on wage payments. With this study, we will contribute to the small but growing body of literature on the effects of audits of public works schemes. Using panel data on 300 panchayats in Andhra Pradesh, Afridi et al. (2014) find that while social audits decreased easy-to-detect corruption related to labor expenditure, harder-to-detect material-related corruption did not decrease, and social audits did not affect the program service delivery. Note that Afridi et al. (2014)'s study was set in the high-performing state of Andhra Pradesh, where the audit processes had been standardized as early as 2006 (Aiyar et al., 2012). We examine the efficacy of social audits through an RCT in a low-capacity setting.[1] Olken (2007) shows that while central auditors reduced corruption in public works programs in Indonesia, publicly visible community-based monitoring could not reduce missing expenditures. This study, on the other hand, evaluates a monitoring system where central auditors encourage and facilitate community participation, and therefore, document the net effect of these two modes of monitoring public service delivery.

Unpacking the mechanisms driving these impacts is essential for understanding how to improve the social audit process in the future. To this end, we will conduct the endline survey in a sub-sample of study gram panchayats (GPs).[2] Specifically, our data and analysis is designed to assess the extent to which declines in work activity in official records are declines in graft (taking the form of listing workers who never worked on program payrolls) or actual declines in work provision to needy households during a time of immense economic distress. The follow-up survey is also designed to assess the extent to which social audits reduce local officials' effort to provide work for local households and the extent to which workers themselves may eschew the program in favor of other work opportunities after an audit takes place (e.g., due to concerns related to corruption or other negative aspects of the program highlighted through the audit.) Therefore, this study will contribute to the literature on the impact of innovations aimed at improving implementation efficiency of MGNREGS (see for instance, Muralidharan et al. (2016)), along with the broader literature on social protection outlined above.

This document summarizes the main tests that we intend to conduct. We do not rule out the possibility of running other tests. We will make clear in the paper which estimations are specified in this pre-analysis plan and which are not (Banerjee et al., 2020).

---

[1]Bihar is one of the poorest regions in the world, and has limited capacity to provide MGNREGS work. Historically, it has struggled to conduct effective social audits (Dutta et al., 2014).

[2]A gram panchayat is a collection of villages and the lowest unit of administration in rural India.

# 2 Background and Experimental Design

## 2.1 Background: MGNREGS Implementation

At the beginning of each fiscal year, local GP leaders prepare a list of potential projects, known as "works", that can be undertaken in the GP by MGNREGS workers. Generally, works are capital improvements, and they are meant to create assets for the local community, such as ponds, roads, or individual assets like latrines. When an individual would like to work for MGNREGS, they must request, or "demand", work from a local MGNREGS implementer. In order to demand and receive work, an individual must be listed on their household's "job card". This document, meant to be issued to all rural households, confirms MGNREGS entitlement eligibility and includes the names of all adult nuclear household members. When there is sufficient work demand, the local official in charge of MGNREGS implementation [3] should open a worksite identified in the list of approved projects. Within a district, materials expenditure can at most be two-thirds of the labor expenditure.

However, MGNREGS is vulnerable to corruption and often faces significant implementation challenges, particularly in low capacity settings. Estimates derived from comparing nationally representative surveys and officially reported numbers of work done under MGNREGS find a leakage rate between 42% and 56% (Imbert and Papp, 2011). Another study in Odisha showed that a wage increase intended for workers was essentially negated through increases in graft by local officials (Niehaus and Sukhtankar, 2013). Leakage in the *wage budget* can occur when officials either pay workers *partial wages* for work completed, with or without the workers' knowledge, or through *ghost work*, where "workers" listed on official program rolls did not actually work. There are two types of ghost work: officials can list individuals who do not exist (less common) or officials list citizens who exist and have a job card, but did not work. In this case, sometimes citizens withdraw wages for work they did not do and share a kickback with local officials.[4]

Beyond corruption, MGNREGS faces a variety of other implementation challenges. For instance, potential workers may not be aware of their right to demand work, be paid an unemployment allowance, or be paid on time. Even when informed, they may lack information about how to exercise these rights. This lack of awareness may, in turn, discourage demand for MGNREGS work. When work is actually provided, workers sometimes must wait a long time to receive wages into their bank accounts, which dilutes the program's value as an emergency source of income.

---

[3]In Bihar, the Panchayat Rozgar Sehak, or PRS.

[4]Leakage can also occur through the *materials budget*, for instance, if local leaders claim more expenses for building materials than they actually make.

## 2.2  Social Audits in MGNREGS

By design, social audits are intended to reduce corruption, improve local accountability, provide officials with inputs to improve service delivery, and give workers recourse to address problems they encounter. Auditors are also responsible for increasing awareness about rights and entitlements under MGNREGS to encourage workers to make use of the workfare program in the future.

Auditors must come from outside the GP and be employees of civil society organizations or other identified and trained individuals, often women affiliated with Bihar's self-help group (SHG) movement, JEEViKA. Employing these women as social auditors presents significant challenges, but is also an important step to empower women as change agents in Bihar's highly patriarchal society.

The audit process typically takes eight days[5] and is comprised of three components: *labor verification*, *asset verification*, and the *public hearing*. Labor verification involves visiting all households that worked under MGNREGS last year according to official data to verify they truly worked. During asset verification, auditors assess whether the physical infrastructure listed in administrative data was completed as described and is of satisfactory quality. These two processes take seven days. On the final day of an audit, findings are presented at a public hearing/community meeting called a "Gram Sabha"; the Gram Sabha is also meant to be a venue for resolving disputes and issues related to MGNREGS.[6]

After the public hearing, the Social Audit Society, the nodal agency that implements social audits in Bihar, shares audit findings with the relevant departments. Other important stakeholders in the audit process include local officials and leaders, since they are supposed to provide all the required documentation for the audits. More importantly, they should be held accountable for any lapses found during the audit process. Based on the audit findings, relevant departments can take disciplinary or punitive action against the offenders.

---

[5]Due to Covid-19 restrictions, audits conducted in the 2020-2021 FY were shortened to only five days.

[6]While we focus on MGNREGS outcomes from and MGNREGS was a primary focus for auditors, other social protection schemes were audited at the same time.

## 2.3 Randomization of Social Audits
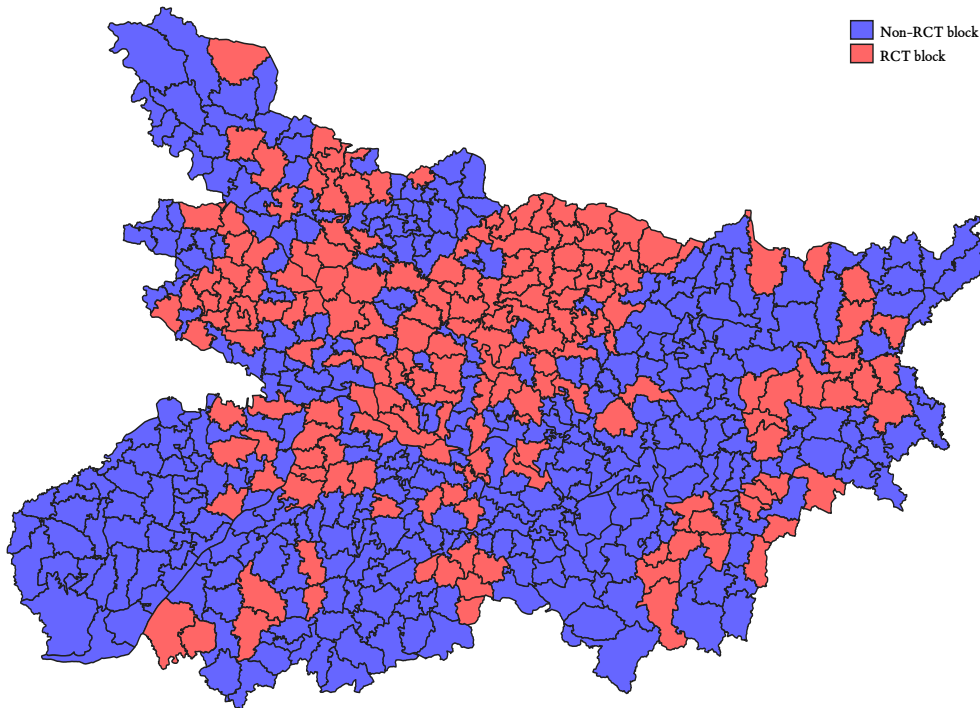
RCT & non-RCT blocks in Bihar



Figure 1: Block level map of Bihar identifying RCT blocks

To identify the causal impact of social audits on MGNREGS program performance, we collaborated with Bihar's Social Audit Society to implement an RCT that leveraged the state's ambitious audit expansion plan in FY 2019-2020. First, the state set targets for the number of GPs it wanted to audit in each block. Then, we randomly selected the GPs to be audited in blocks where less than 100% of locations were targeted for an audit. Thus, our randomization was stratified at the block level, with differing probabilities of treatment driven by SAS's audit target for each block. Figure 1 maps the experimental sample, which includes 3,437 GPs across 178 blocks. Of these, 2,178 GPs were assigned to be audited (treatment), while the remaining 1,259 GPs served as the control group.
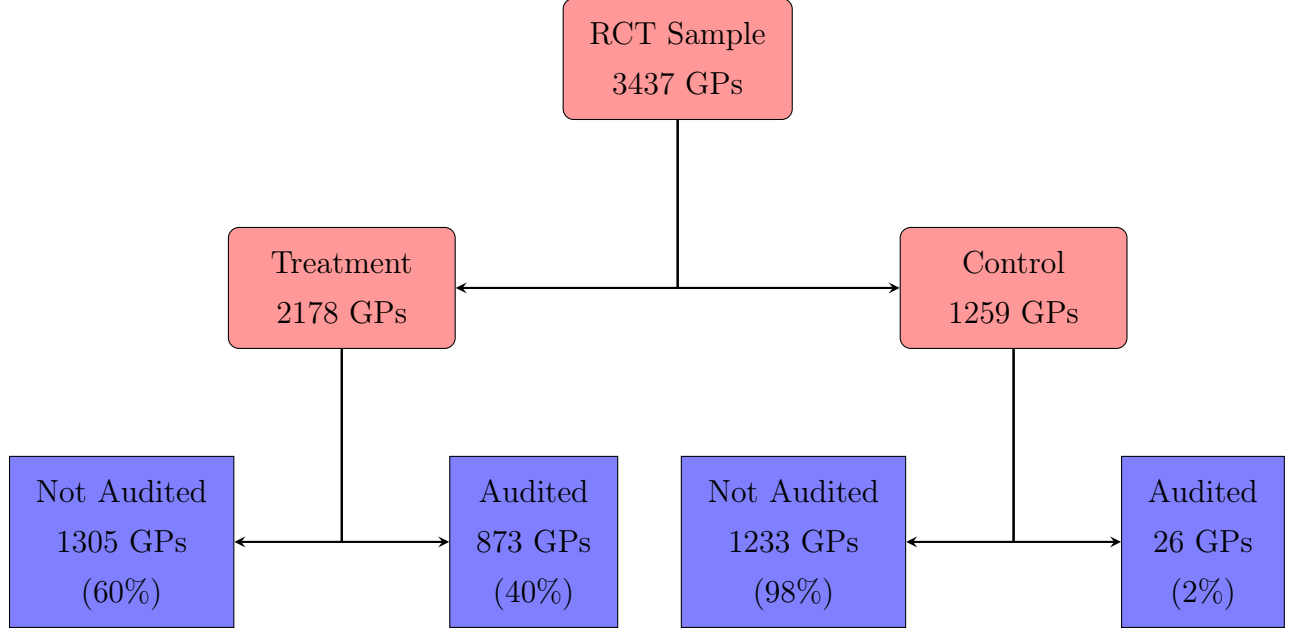
Figure 2: Implementation of Audits in FY 2019-2020

While the state aspired to conduct audits in over 6,000 GPs, ultimately only 40% of the treatment GPs were audited. Figure 2 shows the ultimate treatment status by treatment assignment. One block with randomly assigned audit status was supposed to conducted audits in *all* GPs.

## 2.4 Social Audit Salience Experiment

Since we have shown with administrative data that social audits lead to a decline in wage expenditure, the aim of our endline is to assess whether these reductions reflect reduced corruption, reduced demand for MNGREGS work (e.g. if learning of audit findings discourages workers), or reduced effort on the part of local officials. We will incorporate a salience experiment into the endline to test the discouraged worker hypothesis. In this experiment a randomly selected subset of survey respondents in audit GPs will be given a summary of social audit findings. Respondents treated with this information and "untreated" respondents will be asked whether they would like to work for MGNREGS, and if yes, whether they would like our survey team to share their work request with local officials. This will let us test whether making audit results salient depresses demand for work.

We will only conduct the audit salience experiment within our main RCT treatment GPs, randomizing at the individual level. We will use administrative data to track how many work requests are honored and whether this varies by treatment arm (randomized to audit vs. no audit, randomized to audit salience vs. no salience).

# 3 Research Questions

We aim to identify the causal effect of audits on program outcomes, including program expenditure, work provision, and program leakage. With this objective in mind, in this section we detail our research questions and associated hypotheses.

## 3.1 Primary Outcomes/Research Questions

<u>Research question 1:</u> **How do social audits affect MGNREGS wage expenditures and "demand" for work listed in program records?**
We have already undertaken preliminary examination of administrative data to explore the effect of social audits. Initial analysis suggests that audits reduced the number of workdays paid out by the program, and by extension reduced program wage expenditure. These declines begin when audits start in a local community and persist for months following the audit.

<u>Research question 2:</u> **How do audits affect program graft, measured as (a) the share of official work records that are confirmed by citizen reports and (b) the total amount of work claims that are not confirmed by citizen reports?** As discussed, in the preliminary analysis of administrative data, we note a decline in the program person-days in locations randomly assigned to be audited. In the endline we will ask workers to confirm whether they actually worked at times listed in official program records. This will let us classify offical records in the survey sample into either genuine work or ghost work. We will also record the share of program records linked to individuals who are not residents of the GP or do not exist/were deceased at the time when they were suposed to have worked.

<u>Research question 3:</u> **Do audits affect citizens' reported demand for MGNREGS work?** Social audits may expose fraud or simply the fact that others find it difficult to get leaders to respond to their requests for work on MGNREGS, which could depress demand for work. On the other hand, social audits might raise awareness of the program and increase demand for work. To measure this, we will ask all survey respondents if they would like to for MGNREGS, and if yes whether they would like our survey team to share their work request with local officials. We will also study whether the aforementioned social audit salience experiment affects demand for work in treatment GPs.

<u>Research question 4:</u> **Do audits affect citizens' access to MGNREGS work?** To measure actual receipt of MGNREGS work we will use endline survey responses of workers listed on the MGNREGS rolls. To measure access to work through unofficial channels we will also survey a subset of the general population (drawn from voter rolls) to record the incidence of MGNREGS work among those not found on official program records.

## 3.2 Secondary Outcomes/Research Questions

**Research question 5: Do social audits reduce local officials' willingness to provide work to citizens?** The main way we will answer this question is by triangulating results related to actual receipt of work and demand for work (e.g. if demand is unchanged or increases and actual receipt of work declines, we infer audits reduced local officials' willingness to provide work; if actual receipt of work is unchanged or increases, we infer effort weakly increased). As a secondary check, we will track rates of work receipt among endline survey participants who request that we share their work demand with local officials. This test may be limited in power if demand is low, and further limited by pre-endline turnover among local officials driven by elections and transfers. (The number and composition of citizens who demand work may also be affected by the audit treatment, which complicates this comparison).

Research question 6: How do social audits affect material expenditures relative to labor expenditures?

It has been widely documented that labor expenditure-related misconduct is easier to detect than other more complex expense lines in community-led audits (Afridi et al., 2014; Olken, 2006). Therefore, increasing accountability and punishing offenders in cases of wage misappropriation may have induced officials to misappropriate on the material expenditure margin in the future. To explore this hypothesis, we will examine changes in materials expenditure on MGNREGS projects.

Research question 7: Do social audits increase awareness of entitlements under MGNREGS?

Along with auditing ongoing MGNREGS works and reviewing the scheme's implementation in the past FY, the audit team was also responsible for increasing awareness about MGNREGS entitlements and rules. Moreover, the presence of an external team conducting surveys with MGNREGS participants may increase citizens' awareness about the scheme more generally. Therefore, we will collect data about people's awareness about MGNREGS in the endline survey.

# 4 Data

## 4.1 Administrative Data

To construct measures of program performance, we rely on administrative data posted on the MGNREGS MIS, which we collected via web scrapes. The MIS reports information on monthly employment and demand for work, and we incorporate data from April 2018

through April 2021 in our analysis. To obtain information on jobcards issued and person-days of work disaggregated by caste and gender, we use cumulative totals that are updated daily in the MIS. We scrape these data weekly starting August 2019, and then aggregate the data to the monthly level.[7] Combining both sources of data yields a panel dataset at the GP-month level, which we use in our analysis.

In addition to pulling administrative records of implementation of MGNREGS, we use data on materials and labor expenditure in the study GPs. We scrape these data starting from 2017-18, and similarly construct a GP-month level dataset.

## 4.2   Survey Data

To unpack the results obtained from the administrative data, we plan to conduct a follow-up survey in study areas, which will commence in the first quarter of 2022.

Through these survey data, we aim to measure the impact of social audits on three broad sets of outcomes. First, we want to understand if audits improved awareness of citizens' rights and entitlements under MGNREGS. Therefore, the endline survey will include a module that asks respondents several simple questions about the rules and entitlements of MGNREGS.

Next, we want to measure the impact of audits on demand for work under MGNREGS. As discussed in Section 3, an audit and the subsequent revelation of poor implementation could discourage potential workers from participating in MGNREGS in the future. To measure demand, we will offer to officially register citizens' demand for work with the relevant local official.

In addition to explicitly asking for workers' willingness to register their demand for MGN-REGS, the survey will also include a short module to measure respondents' reservation wage for working on MGNREGS. This section will help establish if workers are discouraged from working under MGNREGS, thereby requiring a higher reservation wage to participate due to the audits.

Finally, we want to ascertain if the incidence of "ghost work" reduced in audited GPs. To do this, we will use publicly available official records on the works conducted in 2019-20 and 2020-21 and use these records to create a sample of MGNREGS workers along with a description of the work they have undertaken in this period. For each sampled worker, we will describe the nature and the timing of work and will record whether they confirm working on the worksite listed by the official records during the relevant time period.

---

[7]To transform and aggregate the weekly cumulative totals to monthly flow measures, we use the *average delta* method. This entails first taking the difference between two successive totals (normalized by the number of weeks elapsed), followed by averaging these weekly deltas over a month. Finally, we multiply this monthly average delta by the number of weeks in the month — to adjust for occasional irregular periodicity of the scrape — to arrive at the desired outcome.

## 4.3 Sampling for Survey Data

### 4.3.1 Sampling GPs

The follow-up survey will be conducted in a sample of RCT GPs, which we selected using an approach that optimizes power and our ability to understand the mechanisms driving the declines in work reports in administrative data. Given this, we focus on blocks (which are also the strata for the main randomization) that maintain high fidelity to treatment assignment and strong results in the administrative data, while maintaining logistical feasibility.

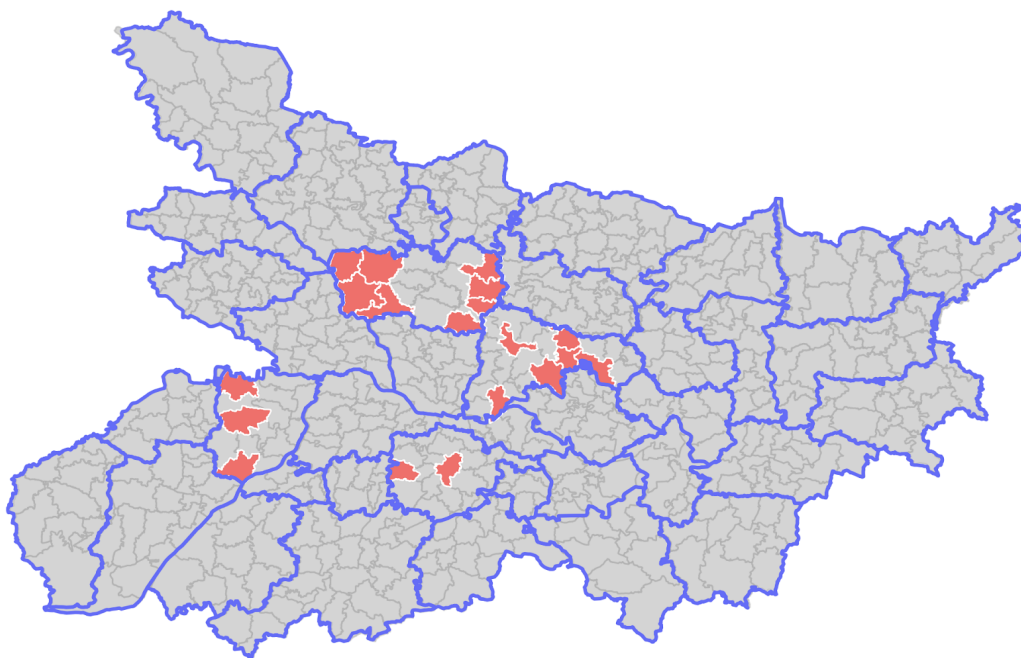**Bhojpur Nalanda Samastipur Muzaffarpur [Fidelity=0.35; 274 GPs]**



Figure 3: Endline sample blocks

We selected the sample as follows. First, we identified a set of districts that were feasible to reach/oversee from our Patna-based field office. Then, we aimed to identify blocks (the randomization strata) that had good adherence to the treatment and a sufficient pool of GPs using the following inclusion criteria:

I. Include blocks that had at least two treatment and two control GPs.

II. Include blocks that had a within-block "first stage" (the share of treatment blocks audited minus the share of control blocks audited) of 0.35 or better

After applying these inclusion criteria we identified 19 blocks from four districts We selected 274 GPs from these blocks as follows. Let $n_b$ for block $b$ the minimum of number of treatment

10

GPs, number of control GPs in the block. We then drew $n_b$ treatment and $n_b$ control GPs from each block, using random selection when the number of GPs in the relevant treatment arm exceed $n_b$.

Using administrative data, we confirmed that the treatment effects are comparable in sign and magnitude with the treatment effects obtained from the full sample. Figure 3 shows the 19 blocks in the sample.

### 4.3.2   Power

To estimate statistical power, we utilized data we collected from a process monitoring survey conducted for social audits in 2019-20 across 13 districts in Bihar, which provides an estimate of the inter-cluster correlation (ICC) for one of the primary outcomes of interest (a dummy variable equal to one if the respondent worked for MGNREGS in the past FY). This ICC is 0.04. With the sample size of 274 GPs available in the 19 selected blocks using the sample selection rules above, and keeping a targeted sample size of 15 respondents listed on administrative data as MGNREGS workers, we reach a standardized minimum detectable effect (SMDE) of 0.109 units. The standardized treatment effect on the person-days worked calculated from administrative data for the chosen sub-sample is very similar (0.09 units).

### 4.3.3   Sampling Respondents

Within each GP, we aim to conduct an endline survey with at least 20 respondents, 15 of whom will be drawn from a publicly available list of MGNREGA workers, which also lists the MGNREGS works undertaken by the listed workers (the "MGNREGS sample"). We will survey five additional respondents from the general adult population. To do this, we will use publicly available voter lists to draw what we call the "non-MGNREGS" sample.[8] Based on piloting, we anticipate that we will be able to locate and interview 75% of individuals listed on MGNREGS rolls and 50% of individuals listed on the voter rolls. Thus we will draw a sample of 20 individuals listed on the MGNREGS rolls and 10 individuals listed on the voter rolls. Tolerating some non-response, we will aim to survey all the respondents listed in the sample. The sample will be stratified by gender within each GP, as will the audit salience treatment.

---

[8]We follow Muralidharan et al. (2016) to decide the ratio (3:1) of MGNREGS and non-MGNREGS samples.

# 5 Empirical Approach

## 5.1 Analysis with Administrative Data

Since all GPs originally assigned to be in the treatment group were not audited, our analysis focuses on "intent-to-treat" effects. Setting aside the possibility of spillover effects, we expect the impact of audits *on GPs that were actually audited* to be a little over twice as large as our main estimates.[9]

We use a Difference-in-Differences (DiD) framework to analyze administrative data results and track treatment effects over time. Within each block we classify the audit cycle for each GP into four mutually exclusive phases:

- *Pre-Audit Phase* — This phase covers all dates prior to April 2019, when the audit calendars were first created. Given our randomization, differences in treatment versus control during this period should be driven by chance.

- *Anticipation Phase* — This phase includes the months leading up to the audit, starting from April 2019 and ending before the first audit took place in a given block. If GPs were notified in advance of their audits, behavior could change in this period.

- *Implementation Phase* — This phase includes the months spanning the first and last audit that took place in the block.

- *Post-Implementation Phase* — This phase includes the months after the last audit took place in the block.

Note that since audits took place at different times for different blocks, the calendar months corresponding to each phase varies across blocks. We interact the treatment indicator with the aforementioned phase indicators to estimate the following regression *in blocks where audits took place*:

$$y_{gmt} = \alpha_{bm} + \beta_1 Treat_g + \beta_2 Treat_g \times Anticipation_{bt} + \beta_3 Treat_g \times Implementation_{bt}$$
$$+ \beta_4 Treat_g \times PostImplementation_{bt} + \varepsilon_{gmt} \tag{5.1}$$

where $y_{gmt}$ represents the outcome of interest $t$ periods from audit in GP $g$ in calendar month $m$. $Treat_t$ is the treatment indicator, which is also interacted with the audit phase indicators

---

[9]Recall that ultimately only 40% of treatment GPs were audited.

as discussed above. $\alpha_{bm}$ is a vector of fixed effects at the block-month level and standard errors are clustered at the block level.[10]

In this regression, $\beta_1$ measures the difference between the treatment and control units before the audit lists were drawn up by the SAS. Relative to this "baseline", $\beta_2$ through $\beta_4$ capture the difference between treatment and control units in each of the phases defined above.

Our preliminary analysis using MGNREGS administrative data indicates that audits had a negative and significant impact on work provision. The average treatment GP records nearly 118 fewer person days of work during the audit implementation period in the block (compared to 1,141 person days of work in the control group during the implementation period). There is no evidence of strategic anticipatory effects that simply displaced work reports to different time periods.

In addition to examining the effect of social audits on key reported administrative data outcomes (work demand, job cards issued, work provision, and share of work provided to women and persons from lower caste groups), we may also examine the following outcomes obtained from administrative data sources:

- Material Expenditure and labor expenditure
- Number of unique job cards against which work was created in a GP
- Days of work per unique job card

## 5.2 Analysis with Survey Data

To fully unpack the mechanisms driving the impact of social audits, we will collect and analyze follow-up survey data. As before, we will focus on "intent-to-treat" effects. We will begin our inquiry by estimating a contemporaneous regression model for the whole survey sample. We present an example below where the the comparison group is the control group:

$$y_{igb} = \beta_0 + \beta_1 Treat_{gb} + \beta_2 female_{igb} + \beta_3 voter_{igb} + \lambda_b + \epsilon_{igb} \tag{5.2}$$

where $y_{igb}$ is outcome $y$ for individual $i$ in GP $g$ in block $b$, $female_{igb}$ identifies women, $voter_{igb}$ identifies individuals drawn from the voter rolls sample, and $\lambda_b$ are block fixed effects. We will cluster standard errors are clustered at the GP level, as this is the unit of randomization.

Our analysis will focus on the following families of outcomes:

Primary Outcomes:
- Work for MGNREGS

---

[10]We do not include "main effects" for the different audit periods because these are absorbed by the block-month fixed effects.

- Demand for work under MGNREGS
- Incidence of ghost work (share of "official" workers who actually worked; value of "official" wages that are genuine). Note that this analysis will be limited to the MGNREGS sample.

Secondary Outcomes:
- Awareness of MGNREGS entitlements
- Experience of misconduct while receiving wage payments (e.g. worker reports of withdrawing funds from their accounts and giving them to local officials)

We will conduct analyses on the pooled sample, as well as the subset of the sample drawn from the MGNREGS rolls, since we expect work among the voter rolls sample to be rare.

## 5.3    Analysis of the Social Audits Salience Experiment

The final element of this study involves understanding if audits affect citizen's willingness to engage with the state and participate in MGNREGS. In addition to analyzing demand for MGNREGS work using the above-detailed approach, we will also analyze results of the social audits salience experiment. This analysis will be restricted to treatment GPs, where we will share the results of social audits with randomly chosen respondents but measure the willingness to engage with MGNREGS for all respondents in treatment GPs.

$$y_{igb} = \beta_0 + \beta_1 Salience_{igb} + \beta_2 female_{igb} + \beta_3 voter_{igb} + \lambda_b + \epsilon_{igb} \tag{5.2}$$

where $y_{igb}$ is the response (for instance, willingness to share demand for work) for individual $i$ in GP $g$ in block $b$. We will use heteroskedasticity robust standard errors for this analysis as the randomization is at the individual level.

We will conduct analyses on the pooled sample, as well as the subset of the sample drawn from the MGNREGS rolls, since we expect demand among the voter rolls sample to be rare.

## 5.4    Heterogeneity

We also plan to study heterogeneous treatment effects to shed light on which sub-populations are the most affected by audits, if any. We expect to explore heterogeneous treatment effects along the following dimensions:

- *Gender*: Here, we will seek to understand if women were more likely to be listed as ghost workers or if women-led social audits were particularly effective in improving women's awareness about their rights and entitlements thereby improving their access

to MGNREGS. As discussed above, we will stratify the sample of respondents by gender such that we have an equal representation of both genders in the intended sample.

- *Caste*: We will examine if respondents belong to marginalized caste groups (scheduled caste or scheduled tribe) are differentially impacted by social audits.

# References

Afridi, F., V. Iversen, et al. (2014). Social audits and MGNREGA delivery: Lessons from Andhra Pradesh. In *India policy forum*, Volume 10, pp. 297–341. National Council of Applied Economic Research.

Aiyar, Y., S. K. Mehta, and S. Samji (2012). India: Implementing social audits.

Banerjee, A., E. Duflo, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautmann (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for rcts in economics. Technical report, National Bureau of Economic Research.

Brown, C., M. Ravallion, and D. Van de Walle (2018). A poor means test? Econometric targeting in Africa. *Journal of Development Economics 134*, 109–124.

Dutta, P., R. Murgai, M. Ravallion, and D. Van de Walle (2014). *Right to work?: Assessing India's employment guarantee scheme in Bihar*. World Bank Publications.

Finan, F., B. A. Olken, and R. Pande (2017). The personnel economics of the developing state. *Handbook of economic field experiments 2*, 467–514.

Imbert, C. and J. Papp (2011). Estimating leakages in India's employment guarantee. *Battle for Employment Guarantee*, 269–78.

Molina, E., L. Carella, A. Pacheco, G. Cruces, and L. Gasparini (2016). Community monitoring interventions to curb corruption and increase access and quality of service delivery in low-and middle-income countries: A systematic review. *Campbell Systematic Reviews 12*(1), 1–204.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2016). Building state capacity: Evidence from biometric smartcards in india. *American Economic Review 106*(10), 2895–2929.

Niehaus, P. and S. Sukhtankar (2013). The marginal rate of corruption in public programs: Evidence from india. *Journal of Public Economics 104*, 52–64.

Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from indonesia. *Journal of Public Economics 90*(4-5), 853–870.

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

Olken, B. A. and R. Pande (2012). Corruption in developing countries. *Annual Review of Economics 4*(1), 479–509.

Ravallion, M. and S. Chen (2008). Does the Di Bao program guarantee a minimum income in China's cities? *Public finance in China*, 317.