Scaling Accountability: Experimental Evidence on Impacts of Citizen-Led Audits on Social Protection Delivery and Election Outcomes*

Charity Troyer Moore
The Ohio State University

Rohini Pande
Yale University

Simone Schaner
University of Southern California

ⓡRandomized author order

February 25, 2025

Abstract

Combating graft while sustaining provider incentives for public service provision in low state capacity settings remains a challenge. We partnered with an Indian state to experimentally evaluate "social audits", one of the world's largest community-driven transparency and accountability initiatives. A social audit engages rural laypeople in week-long monitoring visits, culminating in a public meeting and audit report submitted to the state. Using administrative data from 3,400 communities, surveys and audit reports we find: First, social audits successfully identified program irregularities and engaged communities. Second, these audits affected subsequent implementation of audited programs – rural housing and workfare – despite no change in total expenditure on either program. For housing, where implementers had limited discretion over monitored implementation targets, citizens reported a 28% reduction in kickbacks given to local implementers. Implementers had more discretion in workfare spending, and in treated communities we find a reallocation of program spending on labor payments, where graft is easy to detect through social audits, to materials spending, where graft is harder to detect. Consequently, citizen access to workfare fell by over 50%. Third, social audits changed the nature of returns to elected office, resulting in reduced electoral competition but improved electoral selection: citizens in treated communities are more satisfied with newly-elected local leaders, and subsequent audits identified fewer issues.

*This research was made possible by collaboration and support from Bihar's Department of Rural Development and Social Audit Unit. We are indebted to Bihari officials and citizens who shared their experiences with us. Madhulika Khanna did an exceptional job leading many aspects of the research operation. Akash Bhatt, Geet Chawla, Vidushi Dhawan, Katia Everke, Sameeksha Jain, Shreya Mathur, Vyana McIntyre, Akshita Mehra, Mridul Mishra, Vijay Prakash, Amol Raswan, Bhavya Sinha, Chandan Kumar Sinha, Aruj Shukla, Elena Stacy, Varuni Sureddy, Prema Swarupa Saripaka, and Natalie Theys provided exceptional research support. We acknowledge financial support from the Gates Foundation. Project IRB approval was from IRBs at Yale University and Institute for Financial Management and Research, India; AEA RCT ID AEARCTR-0008640. The AEA confirmation code for author order randomization is sKbD54AYqfM1. All errors are our own.

1 Introduction

Social protection programs now cover more than half of the global population (International Labour Office, 2024). While critical for alleviating poverty and redistributing income in lower-income countries, the programs tend to be worst implemented in areas where they are most needed, both across and within countries (Ravallion, 2019). This is especially true for programs administered through delegated governance structures, in which frontline workers and local elected officials take charge of on-the-ground program implementation.

Although auditing public spending can improve program transparency and implementer accountability (Cuneo et al., 2023), audits may be less effective in detecting irregularities in social protection programs in low-income, low-state capacity settings. First, audits typically require professional evaluators, who are costly to train and deploy on a large scale. Second, uncovering spending irregularities in social protection programs entails, in part, tracking down supposed beneficiaries and assessing whether they received the benefits claimed by the system. Both factors create auditing challenges for lower-capacity governments, especially vis-a-vis programs that service large, hard-to-reach populations. They also limit scope for audit impact, which is greater when audit probability and implementation quality is high (Olken, 2007). And, when it comes to program implementation, audits may *reduce* the amount or quality of local service provision if opportunities for graft provide underpaid and overworked local officials implementation incentives (Acemoglu and Verdier, 2000).

In this paper we evaluate India's "social audits" initiative for major social protection programs, which innovates for scale and feasibility in a low-state capacity setting. During 2023-24, social audits occurred in 210,335 municipalities, covering 78 percent of rural communities and impacting over 700 million Indians, showcasing an unprecedented level of engagement and reach.[1] Social audits deploy lightly trained laypeople (here, rural women affiliated with government-organized collectives) as auditors who spend a week in a rural community verifying program implementation, collecting grievances, informing citizens of their rights and entitlements, and identifying misconduct. At the end of the week, audit findings are shared with citizens and local implementers, who identify action plans to resolve grievances and identified malfeasance. Final reports are shared with the state and made public online; the state monitors issue resolution and has the right to take disciplinary action based on audit findings.

Our evaluation occurred in Bihar, one of India's largest, poorest, and lowest-capacity states. In collaboration with Bihar's Social Audit Unit, we enrolled 40 percent of rural communities in a randomized evaluation of social audits during the state's first at-scale audit rollout in 2019-20. Our analysis covers audits' impacts on two social protection programs. First, the workfare program (the Mahatma Gandhi National Rural Employment Guarantee Scheme, or MGNREGS), which entitles rural households to up to 100 days of minimum wage work. Local program implementers identify public works projects, develop annual workplans and budgets, organize work sites, and procure ma-

[1] By way of comparison, Brazil's well-studied municipal audits program randomly audits roughly 180 of over 5,000 eligible municipalities per year.

terials. Workfare pays a minimum wage; the average study participant received US$24 per ≈10 day work spell. Second, the rural housing program (Pradhan Mantri Awas Yojana-Gramin, or PMAY-G), which provides permanent homes to rural households via conditional cash transfers. Housing beneficiaries are eligible for three fixed size payments totaling US$1,720. Local implementers enroll beneficiaries and certify progress constructing homes, which is required for beneficiaries to access subsequent payments. Both programs use a common electronic funds system to transfer funds to beneficiary bank accounts.

In addition to transfer size differences, the two programs differ in monitoring and the discretion afforded to local implementers. For workfare, implementers are monitored on whether they meet annual budgets (which must allocate at least 60 percent of spending to wages) and project execution. For a given project, implementers have discretion over beneficiary selection and, by extension, the amount allocated to wages versus materials. For housing projects, annual beneficiary targets and lists are distributed by higher authorities to local implementers, who are monitored on progress towards these targets. Discretion to deviate from annual plans is limited. These differing programmatic features impact *how* funds leakage occurs. Leakage under workfare largely occurs through "ghost work": the practice of creating false work records for people who did not actually work. During our study period, beneficiaries in control group communities confirmed only one-third of the work spells recorded in administrative spending data. Leakage can also occur through the materials budget by e.g. over-invoicing. Leakage in the housing program is fundamentally different: while ghost beneficiaries are rare, over 40 percent of beneficiaries in our household survey reported paying a kickback to local implementers to get benefits. The average reported kickback was roughly US$175 (10 percent of program benefits).

Based on our tracking of audits, program implementation, and election outcomes for up to four years after audit roll-out, we report three sets of findings. First, audits successfully identified issues, despite implementation challenges. Second, audits had a deterrence effect, with lasting but different impacts on housing and workfare programs. Graft decreased under housing, where implementers had a narrow set of easily-audited tasks (onboard and approve progress of program beneficiaries) and a single way to appropriate funds. Specifically, while housing program spending was unchanged, treatment-on-the-treated estimates from surveys indicate that beneficiaries were 9 percentage points less likely to pay a kickback, resulting in a 25 percent decline in program leakage. Under workfare, where implementers can appropriate both labor and materials expenditure (the latter is considered harder to monitor (Afridi et al., 2014), and arguably more costly to appropriate), graft appears to have been re-directed towards materials. In particular, the number of workfare projects was unchanged but wage expenditures fell, offset by an (insignificant) increase in expenditure on materials, increasing the materials share of expenditure. Audits had no impact on the *share* of wage payments that were ghost work. Consistent with this, treatment-on-the-treated effects point to a 3.5 percentage point (over 50 percent) reduction in program access among the general population.

Finally, audits had lasting effects, most likely through an electoral channel. Our audit interven-

tion occurred two years before Bihar's 2021 local elections. In audited communities, the number of candidates running for election – who, if elected, would oversee local program implementation, and often benefit from extracting funds during service delivery – declined significantly. Two pieces of evidence suggest audits deterred politicians more motivated by graft, resulting in favorable electoral selection. First, in post-election surveys citizens in treatment areas were more likely to report that local leaders could help them receive community benefits and likewise indicated *increased* interest in participating in workfare, despite no change in the workfare reservation wage. Second, future rounds of audits (uncorrelated with treatment assignment) identified fewer post election issues in treatment communities.

Overall, citizen-led audits reduced spending irregularities, and the combination of top-down monitoring and bottom-up collective action improved accountability. Applying our experimental estimates to all audits conducted during our study period implies that audits reduced housing program leakage by US$10-38 million. We further estimate a US$6.6 million decline in ghost work under workfare, which may have been offset by increased material expense leakage. This comes at a cost of US$3 million less in workfare wages. Bihar paid roughly US$486,000 to conduct audits during the fiscal year we study. Thus, audits can be exceptionally cost effective, with key social planner considerations relating to the relative welfare weights placed on workfare versus housing beneficiaries, and the long-run value of improved political selection. Identifying strategies to mitigate unintended effects on workfare (such as, for example, improved monitoring of the labor versus materials budget) would further strengthen cost effectiveness.

Our work makes novel contributions to the literature on government monitoring and service delivery.[2] The idea that graft might provide implementation incentives goes back to Leff (1964) and Leys (1965).[3] However, there is limited supportive empirical evidence.[4] We provide some of the first well-identified evidence that, when implementers have significant discretion, state-led monitoring can *reduce* service delivery (here, workfare jobs). This reduction occurs alongside increased interest in program participation among citizens, pointing to reduced implementer effort to provide work.

Beyond demonstrating that monitoring can indeed reduce access to benefits, our research underscores that this need not be the case – the *same audit program* implemented in the *same setting* had markedly different impacts on housing and workfare, highlighting the importance of interactions between the design of auditing initiatives and the programs they seek to monitor. In our case, audits only led to reduced citizen benefits when implementers had more latitude to redirect

[2]For recent reviews, see Finan et al. (2017) and J-PAL (2019).

[3]Alternatively, monitoring could reduce implementation effort via multitasking and/or red tape (Banerjee et al., 2012). Given that we find impacts well the audit implementation period, this channel is less likely.

[4]In a notable exception Zamboni and Litschig (2018) find that increased risk of a municipal audit in Brazil reduced graft but had no impact on citizen satisfaction with healthcare services. These outcomes, however, are narrow and not under the immediate control of those most directly targeted by audits – politicians and bureaucrats in charge of public procurement. The literature on community-based monitoring places more focus on outcomes like program access and quality, but results are mixed and these initiatives are typically not focused on rooting out graft, which could be an important determinant of implementer effort (Molina et al., 2016).

graft. These findings enrich a body of work showing that audits conducted by skilled professionals reduce corruption in middle-income settings (Olken, 2007; Bobonis et al., 2016; Avis et al., 2018; Zamboni and Litschig, 2018). Closest to our setting is Wong (2024), who studies the anticipatory effects of social audits in another Indian state. In line with our results, she documents redirection of workfare expenditure from wages to materials, with results pointing to potential increases in misappropriation of those funds. She lacks data, however, to assess impacts on citizens' access to benefits.

Finally, our findings contribute to literature on how audits affect the selection and behavior of elected representatives. Multiple studies document that voters respond to information in "timely" pre-election audits, punishing incumbents with poor reports (Ferraz and Finan, 2008; Bobonis et al., 2016; Larreguy et al., 2020; Berliner and Wehner, 2022). Evidence on the longer-term impact of audits is mixed, with (Bobonis et al., 2016) finding that audits in Puerto Rico do not have lasting effects on graft, arguably because timely audits help citizens select competent, but opportunistic, leaders. Avis et al. (2018), in contrast, find long-term impacts of Brazil's municipal auditing initiative. Their results are not driven by politician selection, but rather deterrence stemming from risk of non-electoral sanction. We advance this literature in two ways. First, we show that audits' electoral effects extend beyond helping voters punish or reward incumbents. Bihar's social audits affected the pool of candidates running for office in the first place (including non-incumbents), highlighting a new dimension of selection that need not rely on voter responsiveness to information. Second, we demonstrate scope for audits to have long-term effects on both service provision and citizen satisfaction, even absent non-electoral sanction of elected leaders.

The rest of the paper proceeds as follows. First, we provide additional background in Section 2. Section 3 describes our experimental design and data, Section 4 our empirical strategy, Section 5 reports our main results, Section 6 discusses mechanisms, and Section 7 concludes.

2 Background

In this section, we provide contextual background on Bihar and India's social audits program. We then describe relevant features of the two audited programs we study, the national workfare program (the Mahatma Gandhi National Rural Employment Guarantee Scheme, or MGNREGS) and the national rural housing scheme (Pradhan Mantri Awas Yojana-Gramin, or PMAY-G), including an overview of specific implementation vulnerabilities.

2.1 Bihar Context

Bihar, with an estimated population of 131 million in 2024, has the lowest income per capita in India – US$736, compared to an India-wide value of US$2,485.[5] One-third of Bihar's population lives in

[5]For context, Bihar's income level looks more similar to low-income countries like Ethiopia (GDP per capita of US$1,294) and the Democratic Republic of Congo (GDP per capita of US$649) than other parts

poverty, with the majority residing in rural areas (Appendix Figure A1, Panel A). Gender norms are conservative, with Bihari women posting one of the lowest rates of labor force participation in the country (31 percent against a national average of 42 percent (PLFS, 2024)). Limited state capacity and governance issues have always plagued Bihar (Mathew and Moore, 2011). These issues have improved in reecnt years, but nevertheless the bureaucracy remains understaffed and overstretched.

2.2 Social Audits

Background. India's social audit program seeks to address social protection program implementation concerns through "top-down" and "bottom-up" monitoring and accountability. After passage of the 2005 National Rural Employment Guarantee Act, all states were mandated to establish an independent Social Audit Unit in 2011. In 2017, the Government of India piloted audits led by women belonging to self-help groups in 11 states, including Bihar.[6] Alongside, the use of social audits was scaled up, with 34 percent of rural municipalities – called gram panchayats, or henceforth "panchayats" – audited in FY 2016-17, 61 percent in FY 2019-20, and 77 percent in FY 2023-24.[7]

Bihar began conducting social audits in 2017, although few were conducted in the first two years. Beginning in FY 2019–20, the state's Social Audit Unit initiated a rapid expansion of audits – increasing tenfold from 248 the year before to 2,645 – with the majority of new citizen auditors recruited from women's self-help groups.[8] Since our experiment, Bihar has continued to increase coverage, auditing 8,023 of 8,380 panchayats in fiscal year 2023–24. New auditors receive 4 days of training, after which they take a competency test. Very few auditors are screened out at this stage, as the state prioritizes basic skills related to literacy, numeracy, and communication.

Audits typically take eight days, with Bihar's FY 2019–20 audits covering four programs: workfare, housing, India's national subsidized food program, and a Bihar-specific sanitation program targeting open defecation. In an effort to minimize capture, auditors cannot hail from panchayats where they work. After introducing themselves to the community, auditors spend the bulk of their time assessing program implementation; for workfare this involves checking that listed workers actually worked and received their wages, verifying that assets were actually built, and identifying any procedural violations or community grievances. The process is similar for housing, with a focus on verifying construction of homes and beneficiaries' receipt of funds. On the final day, findings are presented at a public hearing called a *gram sabha*; the gram sabha also serves as a venue for

of India. Country-wide figures for GDP per capita are for 2023, retrieved from the World Bank Databank. Bihar-specific estimates are for fiscal year 2023-24 from India's Ministry of Statistics and Programme Implementation.

[6] A self-help group is a government-organized collective of 10–25 women who participate in group savings and loan activities. The government uses self-help groups as a platform for social education and often engages members in a range of socially-oriented ventures.

[7] The gram panchayat is the lowest unit of governance in rural India, typically consisting of roughly 5 villages in Bihar.

[8] The number of completed audits reported online has changed over time as the government updates and refreshes its records. We report the number completed per the main web scrape that we use in our analysis.

resolving citizen concerns related to program implementation. Gram sabha attendees include local implementers, panchayat leaders and citizens, and representatives from higher levels of government.

After the public hearing, the Social Audit Unit shares audit findings with relevant departments. Issues are also discussed at a public hearing held at the block level, an administrative unit covering roughly 16 panchayats. Audit issues are then assigned to a responsible officer at the block or district level, who is tasked with ensuring the issue is resolved. Audit findings and progress towards issue resolution are digitized and tracked by higher-level authorities.[9] Program implementers holding elected office can be held accountable for audit findings by citizens, who are encouraged to attend the community meeting. Unelected implementers can be sanctioned by higher levels of government. The most severe forms of formal sanction including fine, censure, suspension, and termination are (for workfare) posted publicly online. Less severe (and untracked) forms of punishment include warnings, undesirable transfers, or failure to renew/offer new postings to contracted employees.

Audit Implementation in Bihar. In order to understand implementation quality, we collected process monitoring data from 24 panchayats in 13 districts during Bihar's FY 2019-20 audit rollout.[10] During process monitoring visits, a team of trained enumerators shadowed auditors to assess adherence to protocols, interviewed auditors, spoke with beneficiaries slated for an audit visit, and (in one panchayat per district) observed the closing community meeting (gram sabha). Table 1 reports findings from these visits. Panel A reports characteristics of auditors themselves. Over 80 percent were women, 37 percent had less than a 12th grade education, 39 percent were high school graduates, and on average they were 31 years old. Auditors reported having to travel an average of 3 hours to reach the audited panchayat. While all reported receiving training from the state, 61 percent reported that they would like more training. Finally, while just 4 percent of women reported that they felt unsafe doing audit work, 59 percent reported feeling "somewhat safe". This is not entirely surprising, as gender norms in Bihar typically restrict women's mobility and interaction with strangers.

Panel B reports results from enumerator surveys of a random sample of workfare beneficiaries listed on auditors' work verification lists. Only 46 percent reported participating in the workfare program during the audited period, while 25 percent reported receiving payment from the program without participating, highlighting one way implementers siphon funds from the program: asking beneficiaries to withdraw unearned wages in return for a modest payment. Sixty-one percent reported they were interested in getting workfare, but only 14 percent reported ever attempting to demand work from local implementers. We also find evidence of reasonably high audit coverage: 62 percent were aware of the audit team in their community and 46 percent reported receiving a visit from an auditor. Advertisement of the community meeting was less successful: 30 percent of

[9]Audit and action taken reports are publicly available for workfare, but not housing. However, a similar process for tracking actions taken to resolve audit issues occurs for both programs.

[10]The districts were Banka, Begusarai, Bhagalpur, Bhojpur, Gopalganj, Jehanabad, Madhepura, Madhubani, Purbi Champaran, Purnia, Samastipur, Sheohar, and Supal.

respondents we interviewed before the meeting were aware of it, while 15 percent of respondents interviewed after the meeting reported attending. This suggests half of those aware chose to attend, indicating high interest.

Panel C summarizes findings from shadowing of workfare labor verification. While auditors almost always asked listed beneficiaries if they had worked under the program, detailed follow up on was less consistent, with 27 percent asking about the number of days worked, 44 percent asking if wages were sent to a bank account, and 78 percent asking if the household had a job card (required to grant access to the program). Just 6 percent of auditors asked respondents if they would like to participate in workfare, and only 37 percent mentioned the community meeting.

Panel D reports on the community meetings themselves. Overall, engagement was high: community meetings happened in all selected panchayats, with attendance topping 50 people 85 percent of the time and 100 people 46 percent of the time. The elected leader of the panchayat (the mukhiya) attended 92 percent of the meetings we observed, as did the local workfare implementer (the PRS). District-level officials attended over half the meetings. On average, 3.9 citizens gave public testimony of grievances under workfare, while 2.4 citizens testified to grievances related to housing. Our trained observers felt that the mukhiya exerted undue influence on the proceedings 38 percent of the time.

Publicly available records from "action taken reports" indicate that the state took at least some formal steps to follow up on audit findings. Appendix Figure A2, Panel A graphs the total number of workfare issues identified through social audits, breaking them up between those marked as resolved (1,674) and unresolved (15,144). Panel B shows that audits identified over ₹105 million ($1.4 million) in financial misappropriation, of which officials were only able to recover ₹5 million ($60,000). Panel C shows that the state was able to resolve a greater share of wages due to workers.

2.3 Workfare

Background. The Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) is India's second-largest social protection scheme (behind subsidized food) and the largest workfare program in the world; in fiscal year 2020–2021, for example, 75.5 million households were listed as working at a cost of US$13.8 billion. As its name suggests, MGNREGS employs employs rural citizens to work on program-linked projects. These projects are typically capital improvements meant to create assets for the local community, such as ponds or roads. The scheme also supports construction of individual assets like latrines, sometimes through "convergence" with other government programs like India's national sanitation initiative. Although a federal program, workfare is administered and implemented at the panchayat level. The decentralized design aims to empower local communities to participate in workfare-related decision making, both in terms of the type of projects and the amount of work undertaken (Fischer and Ali, 2019).

At the beginning of each fiscal year, local panchayat leaders prepare a list of potential program projects, which informs their annual budget, of which 60 percent must go to worker wages. The

program is meant to be demand-driven, meaning that when an individual would like to participate in workfare, he or she should request work from the government-employed local implementer known as the *Panchayat Rozgar Sewak* (PRS). When there is sufficient work demand, the PRS should open a worksite identified in the list of approved projects. *Mukhiyas* – the elected leaders of panchayats – also play a significant role in implementing workfare, especially in allocating jobs to interested individuals. Workers are paid a state-specific daily wage for each day of work.[11] Wage payments are sent electronically to worker-owned bank accounts. Local implementers are regularly monitored on their progress against annual plans, meaning there is significant pressure to meet an expenditure floor, but less pressure to spend beyond that even when community demand is high. Implementers are monitored on the 60 percent labor target at the district level, leaving some wiggle room for panchayats to deviate.

Workfare Implementation in Bihar. Within India, Bihar consistently ranks at the bottom of Indian states in MGNREGS expenditure per rural poor citizen (see Appendix Figure A1, Panel B). Like other programs, workfare faces important implementation challenges and is vulnerable to leakage, particularly in settings with limited state capacity like Bihar. There is no robust way for higher levels of government to track work demand; thus local implementers have significant discretion regarding when and to whom work is offered. As a result, on ground the program is typically supply-, rather than demand-, driven: work is made available and allocated to citizens by local implementers – namely the PRS and local elected leaders – with notable rationing, meaning many people who would like work never receive it (Ravallion, 2019).

There are two main ways leakage of wage expenditures can occur. First (and most common), implementers may make work claims for individuals who did not actually work, which we refer to as *ghost work*. Since workfare payments are made to worker-owned bank accounts, ghost work either requires the involvement of the listed beneficiary (who may or may not retain a portion of the payment in return for cooperation), or collusion with actors like local bank kiosk operators. Second, individuals who actually worked may be asked to give a kickback to local officials in return for access to the program. Figure 1 uses data from our control group to graph rates of ghost work and kickbacks among workfare beneficiaries (see Section 3.2 for more detail on these data). Panel A shows that ghost work is far more common than kickbacks, with over two thirds of the work spells we crosschecked unverified by beneficiaries. This is in line with both our process monitoring findings and Imbert and Papp (2011)'s estimate that 69-74 percent of MGNREGS expenditure in Bihar is ghost work. Kickbacks are quite rare, affecting less than 4 percent of work spells. The second graph shows that the typical paid beneficiary reports receiving ₹1,980 (US$24 using an exchange rate of ₹82 per $US) per work spell. Beneficiaries asked to provide kickbacks receive less, with more than

[11]The program employs engineers to assess work completion, which is recorded and published online. Partial work can lead to partial payment (which is common in some Indian states), but in Bihar the vast majority of workers are paid the full daily wage (in FYs 2019–20 and 2020–21, the average wage paid was the same as the state's official daily wage).

half of reported payments channeled to local implementers.

Leakage can also occur through *materials expenditures*, e.g. if local leaders over-claim expenses on construction materials, or take kickbacks from materials suppliers in return for patronage. Materials payments are logged in the MGNREGS management information system (MIS) and paid directly to approved materials providers. These payments are slow to arrive (often delayed by more than a year in Bihar); this may be one reason why, per qualitative evidence, local implementers often siphon some worker wage payments (which typically arrive within 15 days of a completed work period) and use these funds to compensate contractors to complete projects.

2.4 Rural Housing

Background. Pradhan Mantri Awas Yojana-Gramin (PMAY-G) aims to construct permanent homes (i.e. cement floors, brick walls) for rural Indian households living in temporary structures (i.e. dirt floors, mud walls). At the time of our study, PMAY-G was one of India's largest social protection programs: in FY 2020–21 the government spent $5.6 billion on construction of 3.4 million homes, almost 1 million of which were in Bihar. The main payments under the program are meant to cover building materials, and are paid to households in three tranches via beneficiary-owned bank accounts, totaling either ₹120,000 or ₹130,000 (US$1,463 or US$1,585). Households are responsible for overseeing construction of the homes on their own, and receive 90-95 days of wage payments under MGNREGS (also delivered in three tranches) to cover construction as part of MGNREGS-PMAY-G "convergence". These labor payments are made regardless of whether households use their own members or hire laborers to construct their house. To ensure compliance with program rules, photos verifying progress on the house must be geotagged and uploaded on the PMAY-G web portal in order to unlock access to the subsequent installments. Local officials —contracted workers called *Awaas Sahayaks*, or MGNREGS implementers, or local leaders —facilitate this process.

Eligibility for PMAY-G is based on housing characteristics recorded in the Socio-Economic and Caste Census, a large-scale survey designed to identify below-poverty-line households, last conducted in 2011 and (for Bihar) a supplementary survey conducted during FY 2018-19. Preliminary lists based on survey data were updated by panchayat officials during a community meeting; once fixed this serves as a wait list. State and district officials give panchayats an annual target of homes to complete, with local implementers monitored on progress against this target. Thus, compared to workfare, local implementers have significantly less discretion over implementation intensity and program expenditure.

Housing Implementation in Bihar. In contrast to its relatively limited workfare activity, Bihar was a leading implementer of the housing program during our study period, with particularly high activity in FY 2019–20 (Appendix Figure A1, Panel C). Like workfare, there are two primary forms of leakage under the housing scheme: first *ghost houses*, which are relatively rare given the geotagging system, and second *kickbacks*, which are more common since local implementers can

easily gatekeep access to subsequent installments via the certification process. Figure 1, Panel B explores leakage under PMAY-G in our control group. Ghost houses are rare – only 14 percent of listed beneficiaries claimed they had not received any funds from the program – but kickbacks are common, reported by 43 percent of beneficiaries. The second graph in Panel B shows that while kickbacks are small as a percent of reported housing payments (which average ₹111,000), they are large in magnitude. The average listed beneficiary reporting a kickback paid ₹14,476 (US$177) – over 7 times the average workfare payment and 24 percent of state per capita GDP.

Program design likely determines the different patterns of corruption under workfare and housing: PMAY-G houses must be photographed, geotagged, and uploaded to an online portal – this makes ghost houses relatively easy to detect, while kickbacks occur at the beneficiary level and are harder to track. For workfare, on the other hand, there was no way to systematically identify ghost work during our study period; and kickbacks may have triggered significant beneficiary resistance given the modest value of the average work spell payment. Thus, both patterns of leakage are consistent with implementers choosing the path most likely to evade detection and minimize resistance among concerned parties.

3 Experimental Design and Data

3.1 Randomization

We collaborated with Bihar's Social Audit Unit to implement an RCT alongside the state's audit expansion in FY 2019–20 (see Appendix Figure A3 for a timeline). First, the state set targets for the number of panchayats it wanted to audit in each block, the administrative unit between a panchayat and a district.[12] Then we randomly selected the panchayats to be audited in blocks where less than 100 percent of panchayats were targeted for an audit. Thus, we randomized at the panchayat level, stratifying on the block, with differing treatment probabilities determined by the state's audit target for each block. Appendix Figure A4 maps the experimental sample, which includes 3,437 panchayats across 178 blocks. Of these, 2,178 panchayats were assigned to be audited (treatment), while the remaining 1,259 panchayats served as the control group.

The state aspired to conduct audits in over 6,000 panchayats, but only 40 percent of treatment panchayats were ultimately audited, reflecting a mix of delays and a sudden halt in auditing due to the Covid-19 pandemic. To maximize power, our main analysis limits attention to the 60 percent of blocks (randomization strata) in the experimental sample where at least one audit happened. Within this sample of 2,101 panchayats, our main specification estimates intent-to-treat effects.[13]

[12]There are an average of 16 panchayats per block, an average of 14 blocks per district, and 38 districts in Bihar.

[13]We find no evidence that audits had an effect in blocks where audits never happened, consistent with the fact that the master social audit schedule was never shared with local officials – rather, panchayats were notified of impending audits shortly before they took place.

Appendix Table A1 verifies randomization balance among panchayats in blocks where audits happened. The first column reports average outcomes in the control group and column 2 reports regression-adjusted differences between the treatment and control group, conditional on block fixed effects. Panel A covers outcomes from administrative data, including measures of wage and materials expenditure under workfare and housing. All outcomes are measured at the panchayat-month level. Panel B reports panchayat-level socioeconomic characteristics from the 2011 Socioeconomic and Caste Census (SECC) including total population, gender composition, and socio-economic status. Overall, the randomization is well-balanced with only one of 12 characteristics significant at conventional levels: treatment panchayats posted slightly lower (5 percent of the control mean) housing material benefits expenditure in the pre-intervention period.

Appendix Table A2 verifies balance among panchayats in blocks where no audits happened (columns 1-3) and explores differences between audited and non-audited panchayats in non-RCT blocks (columns 4-6).[14] Comparing columns 1-3 to 4-6 shows that when given discretion Bihar targeted audits to panchayats with more MGNREGS expenditure. SECC data show that (non-randomly) audited panchayats are slightly smaller and better educated, though generally quite similar to unaudited panchayats in terms of demographics.

3.2 Data

Our analysis combines multiple administrative datasets with original survey data to paint a holistic picture of audits' impacts on program, citizen, and electoral outcomes. We provide an overview here and give a detailed accounting of our data and variable construction in Appendix B.

3.2.1 Administrative Data

To construct measures of program performance, we rely on administrative data from the MGNREGS management information system (MIS), which we collected through web scrapes and regular API-based data pulls, as well as the PMAY-G MIS, also collected through web scrapes.

Measuring MGNREGS Labor and Materials Expenditure. We leverage publicly-available MIS reports to measure panchayat-specific person-days of demand registered, person-days of employment provided (which we map onto wage expenditure by multiplying by the relevant wage rate in a given fiscal year), and materials expenditure. The advantage of these data is that we can use them to build a monthly panel spanning April 2018 (one year before our intervention began) to April 2021 (one year after the last round of RCT-based audits were conducted). The resulting measure of wage expenditure, however, combines payments linked to both traditional workfare and housing.

[14] Data on PMAY-G administrative outcomes are not available for these sub-samples.

Separating Housing- and Workfare-Linked Labor Payments Under MGNREGS. In order to separately study wage payments linked to workfare and housing, we utilize data drawn from an API. This lets us obtain more granular payment information that details the type of work performed. We used this information to build a monthly panel of panchayat-level workfare- and housing-linked labor payments spanning April 2019 to April 2021. The disadvantage of these data is that they are not available for the pre-randomization period.

Measuring Housing Materials Expenditure. We use information from the PMAY-G MIS to construct a panchayat-level monthly panel of newly registered materials expenditure (associated with newly-approved homes) and incurred materials expenditure (a mix of first payments for newly-approved homes and later payments for previously-approved homes). These data are available from April 2018 to April 2021.

Social Audit Reports. To understand audit findings, we scraped social audit reports from the MGNREGS MIS. The reports contain the name of auditors, types of issues identified, community meeting attendance, and (when relevant) amounts of financial misappropriation associated with the issues.

Electoral Outcomes. In order to study audits' effects on electoral outcomes, we digitized data from Bihar's State Election Commission. Elections for panchayat representatives take place on a five year cycle; Bihar went to the polls in 2016 and 2021. We use data on election winners in 2016 to identify incumbents. We link these data (using names) to 2021 election candidates and voting results, which lets us identify whether an incumbent ran for re-election, the number of challengers, who won, and vote shares of incumbents and challengers.

3.2.2 Survey Data

While administrative data are informative of at-scale impacts on program expenditure and payments sent to citizens, they confound genuine payments with leakage. In order to differentiate the two, we conducted an in-person follow-up survey in 274 panchayats in three districts in central Bihar (Bhojpur, Muzaffarpur, and Samastipur).[15] We fielded the survey as soon as it was possible to conduct wide-scale in-person field work post-Covid. The survey took place between March–June 2022, approximately six months after local elections.

Our survey included two sampling frames. First, to map beneficiary experiences under workfare and housing, we web scraped details from beneficiary payments under MGNREGS in our sample panchayats. We used these data to randomly sample 20 listed beneficiaries (10 male, 10 female – including both workfare and housing-linked beneficiaries) per panchayat. Workfare beneficiaries

[15]Note that our pre-analysis plan mentions a fourth district, Nalanda; while our sampling code considered blocks in this district, none were enrolled based on our sampling algorithm, which is detailed in Appendix B.

were asked to confirm participation in up to four workfare "work engagements" during our study window.[16] The survey asked workers about total wage payments received, kickbacks paid to local officials, and the amount retained per work engagement. Housing beneficiaries were asked to report whether they received funds for a PMAY-G house, and if so, the total amount of payments received, as well as kickbacks and the amount retained.[17] The survey also recorded demand for workfare and knowledge about the program. The survey asked about benefits received between April 2019 and March 2021. Thus, one concern is recall bias, especially for workfare, which is less salient than building a new home – to the extent that respondents failed to remember past workfare engagements, this would bias estimated levels of graft up. To mitigate this risk, enumerators prompted respondents listed on the workfare rolls with descriptions of the type of work as described in the MIS data. While we received qualitative reports from the field that that workfare was salient to respondents, our measures of program leakage should nevertheless be interpreted as an upper bound.

In order to measure the general population's access to and demand for workfare we additionally sampled 10 individuals per panchayat (5 male, 5 female) from voter rolls. We screened out extremely wealthy respondents prior to enrolling in the survey, leaving this sample representative of the general population minus local elites. This survey asked respondents whether they had participated in workfare during our reference period and recorded demand for and knowledge of the program.

Our source datasets for both samples had limited information to facilitate tracking respondents. Moreover, many individuals registered to vote in Bihar live out of state due to high rates of migration. In practice, we tracked and interviewed 78 percent of individuals sampled from MGNREGS administrative data (workfare and housing beneficiaries) and 58 percent of individuals sampled from voter rolls. This tracking rate is balanced across treatment groups (see Appendix Table A3).

Appendix Table A4 verifies that demographic characteristics of those surveyed in the voter roll sample are balanced across treatment and control groups. Overall, sampled individuals are middle aged, married, live with 5 other people, and have limited human capital: 45 percent have no education at all and another 30 percent only have primary school education. Nineteen percent come from marginalized communities called scheduled castes or tribes. Only 6 percent are "socially connected" meaning they are related to someone holding a leadership position in the panchayat. While characteristics in the beneficiary sample need not be balanced (since treatment could affect who implementers select for benefits, especially under workfare), Appendix Table A5 shows that there are limited differences across treatment and control.

[16] A work engagement is a spell of continuous work on a specific project, allowing for up to a two week break in work. Roughly 90 percent of individuals had 4 or fewer work engagements, and 86% of work engagements were highly differentiated by type of work such that it was unlikely a respondent would confuse them. See Appendix B for more detail.

[17] We did not ask respondents to differentiate between payments for labor made through MGNREGS convergence and the larger payments for housing materials, since our piloting found limited ability to separate the two.

4 Empirical Strategy

Since we have many months of administrative data, we use difference-in-differences and ANCOVA specifications to estimate treatment effects over time. We classify the audit cycle into three mutually exclusive phases:

- *Pre-Audit Period* — This phase covers all dates prior to April 2019, when the randomization took place and audit calendars were first created. Given our randomization, differences in treatment versus control during this period are driven by chance.
- *Anticipation Period* — This phase includes April and May 2019. The first audits took place in June 2019; since the audits calendar was not publicly posted and the Social Audit Unit did not give local officials much advance notice of audits, we expect to see little-to-no impact in this period.
- *Post Period* — This phase includes the 10 months when audits were taking place (June 2019–March 2020) and 13 months after study-linked audits ended (April 2020–April 2021).

Our baseline difference-in-differences specification is:

$$y_{pt} = \alpha_{bt} + \beta_1 Treat_p + \beta_2 Treat_p \times Anticip_t + \beta_3 Treat_p \times Post_t + \varepsilon_{pt} \quad (1)$$

where y_{pt} represents the outcome of interest in panchayat p in month t. $Treat_p$ is the treatment indicator, which is also interacted with the audit period indicators as discussed above. α_{bt} is a vector of block-month fixed effects. Standard errors are clustered at the block level.[18]

In this regression, β_1 measures the difference between treatment and control panchayats before the audit lists were drawn up by the Social Audit Unit. Relative to this "baseline", β_2 and β_3 capture intent to treat effects in the anticipation and post-treatment periods respectively. Results are robust to defining "block specific" anticipation and post-treatment periods based on the first and last audit dates in each block (see Appendix Table A7).

When using survey data, we run individual-level intent-to-treat analyses using the following specification:

$$y_i = \beta_0 + \beta_1 Treat_p + \beta_2 Female_i + \beta_3 Voter_i + \beta_4 housing_i + \lambda_b + \epsilon_{ibp} \quad (2)$$

where y_i is outcome y for individual i, $Female_i$ identifies women, $Voter_i$ identifies individuals drawn from the voter rolls sample, $housing_i$ identifies housing beneficiaries in the MGNREGS work spell sample, and λ_b are block fixed effects. Standard errors are clustered at the panchayat level.

[18] As expected given the randomization design, results are essentially unchanged when clustering standard errors at the panchayat level.

We drafted a pre-analysis plan for this project prior to the post-intervention follow up (endline) survey, but after observing initial treatment effects in the administrative data.[19] Our pre-specified outcomes of interest are (primary:) administrative expenditures and recorded labor demand, leakage (share of benefits that are fraudulent, expenditure amount not verified by citizens), stated demand for workfare, access to workfare; (secondary:) materials expenditure, workfare awareness, and local officials' willingness to provide workfare.

The above regression specifications conform to the pre-analysis plan, with two exceptions. In the administrative data analysis, we pool audit implementation and post-implementation periods into a single post-treatment period. This is to simplify the analysis given that we observed very similar impacts in these periods (see Appendix Table A8). In the survey data analysis, we add a housing beneficiary dummy to the survey data specification. Since housing benefits are much larger than workfare benefits, this improves precision in specifications that pool across beneficiary types.[20]

Since some of the administrative data that we collected post pre-analysis plan is only available in the post-intervention period, we also use ANCOVA and "modified" ANCOVA specifications in our administrative data analysis. When pre-period data is available, the ANCOVA specification drops $Treat_p$ in equation 1 and controls for the average outcome from April 2018–March 2019. For outcomes where we lack pre-period data, we use a "modified" ANCOVA specification that controls for pre-period averages of registered labor demand and expenditure from the MGNREGS MIS and housing registrations and materials benefits expenditure from the PMAY-G MIS.

5 Main Results

We begin by documenting our first stage, followed by a significant decline in labor expenditure, driven by a reduction in workfare (rather than housing) wage payments. We then assess the extent to which these changes reflect declines in program leakage versus changes in program access and compare the cost of program implementation to audits' effects on these outcomes.

5.1 First Stage

Appendix Table A6 shows the first stage for the full RCT sample (limited to blocks where audits happened), followed by the first stage for the follow up survey sample. Overall fidelity to the randomization was very good – just 3 percent of control panchayats received an audit, and the likelihood of getting audited in FY2019-20 was 66 and 72 percentage points higher in the full RCT and survey samples respectively. Columns 2 and 5 show that treatment had no effect on the state's

[19]The pre-analysis plan is available at `https://www.socialscienceregistry.org/trials/8640`, RCT ID AEARCTR-0008640.

[20]The pre-analysis plan focuses on MGNREGS as a whole and does not separately discuss PMAY-G. We decided to separately focus on PMAY-G after completing sampling for the endline survey, which highlighted significant PMAY-MGNREGS convergence.

decision-making over which panchayats to audit in future fiscal years. As a result, at the end of FY 2023-24, our treatment induced a 32-33 percent increase in cumulative audit exposure (columns 3 and 6).

5.2 Impacts on Administrative Outcomes

Table 2 studies impacts on administrative outlays for citizen-facing benefits under workfare and housing. Panel A of reports impacts on labor benefits made through MGNREGS. Demand registration indicates implementers' efforts to initiate workfare projects, as officially logging requests for work in the MIS is one of the first steps to provide work and rarely done to simply record interest in the program. Wage expenditure summarizes workfare- and housing-linked labor benefits officially sent to citizens. Overall, we find no impacts during the anticipation period (as expected) and evidence of modest, significant declines in both registration and wage expenditure in the post period. Effects are on the order of 4 percent of the dependent variable pre-period mean – or 6 percent when we instrument audit receipt with treatment to calculate treatment-on-the-treated effects (Appendix Table A9).[21] Panel B shows that audits had no impact on housing registrations (an analog to demand registration, in that this action is required to initiate payments under the program) and expenditure on housing materials benefits. These point estimates are small in magnitude (1.2 percent of the dependent variable mean or less) and precise – 95 percent confidence intervals rule out declines larger than 3.6 percent of the dependent variable mean.

The estimates in Table 2, Panel A are for total MGNREGS wage expenditure, which combines workfare and housing outlays. While the null effects on housing materials expenditure in Panel B) hint that these declines reflect reduced expenditure on workfare, this need not be the case, as different functionaries and administrative systems are responsible for processing housing materials and labor payments. Panel A of Table 3 assesses treatment effects using our alternate MGNREGS data, which – while not available for the pre-intervention period – distinguishes between workfare- and housing-linked wage expenditure. Column 1 considers total wage expenditure; these estimates are in line with those in Table 2, column 6.[22] Columns 2 and 3 decompose impacts into expenditure associated with workfare and housing respectively. Effects are almost entirely driven by a reduction in workfare (a 5 percent reduction in the dependent variable mean, accounting for 80 percent of the total effect size). Impacts on housing outlays are small in magnitude (2 percent of the dependent variable mean) and not statistically significant.

In addition to the beneficiary-facing outlays discussed so far, a significant amount of the workfare

[21] Insofar as treatment had spillover effects on neighboring panchayats treatment-on-the-treated estimates will be downward biased. Consistent with no anticipatory effects, Appendix Table A10 shows that audits had no effects in the subset of RCT blocks where audits never happened. This table does not include housing registrations and materials payments because these data are only available for blocks where audits happened.

[22] The two data sources are drawn from different parts of the MGNREGS MIS and panchayat-level averages, while highly correlated ($rho = 0.93$), do not exactly match. Therefore our estimates are slightly different.

budget goes towards materials expenditures (recall that at least 60 percent of expenditure must go to labor, leaving up to 40 percent for materials). Table 3, column 4 reports impacts on workfare material expenses. While point estimates are not significant, they are similar in magnitude to the treatment effects on wage expenditures in column 1 and in Table 2. Thus on net, audits had limited impact on total expenditure falling under MGNREGS. Column 5 of Table 3 shows that audits reduced the wage share of all MGNREGS expenditures by 1.3 percentage points, while column 6 shows that the wage share of workfare-specific expenditures (setting aside housing-linked labor benefits) fell by 3 percentage points.[23] This does not reflect an increase in asset building: audits had no impact on the total number of projects executed under MGNREGS (Appendix Table A12) and point estimates are very small (less than 1 percent) relative to dependent variable means. The shift in the wage expenditure share could reflect better adherence to program rules if leaders stopped relying on ghost work to finance payments to materials vendors. However it could also reflect a reallocation of – or even an increase in – graft if funds diverted to materials payments were used inappropriately. We view this as more likely, since social auditors are better placed to detect ghost work than misuse of materials payments (Afridi et al., 2014).

5.3 Impacts on Misallocation and Program Implementation

The reductions in workfare wage expenditure could reflect less graft (possibly reallocated to materials expenditure) or fewer beneficiary payments, with the latter stemming from either lower levels of demand among citizens or lower effort among implementers. Table 4 uses the follow up survey's beneficiary sample to assess intent-to-treat effects on graft *among official program beneficiaries*; Appendix Table A13 reports treatment-on-the-treated effects. Panel A pools observations associated with workfare and housing beneficiaries together to paint an overall picture of impact. Since patterns of leakage are very different in the two programs (recall Figure 1), Panel B decomposes results by beneficiary type.[24] The first column reports impacts on an overall "misallocation index", which is a standardized, GLS-weighted average of the outcomes in columns 2-6, following (Anderson, 2008). Overall, audits reduced misallocation by 0.09 standard deviation units, largely through reduced kickbacks (column 4) and – as a result – reduced leakage (column 6), which we define as the difference between administrative outlays and payments respondents reported keeping (column 5).

[23]The total wage expenditure share divides total wage expenditure – including both workfare- and housing-linked payments by the sum of total wage expenditure and workfare materials. The workfare wage expenditure share divides workfare-linked wage expenditures by the sum of workfare-linked wage expenditures and workfare material expenditures. Since workfare materials expenditures are invoiced sporadically and expenditure shares tracked by fiscal year, we calculate wage expenditure shares by aggregating across the entire anticipation period (April and May 2019), the implementation period (which corresponds to the remainder of FY 2019-20), and the post-implementation period (which corresponds to FY 2020-21, plus April 2021). These regressions include block×aggregation period fixed effects instead of block×month fixed effects.

[24]127 survey respondents received both workfare- and housing-linked benefits; we stack the data so there is one observation for each program.

Panel B shows that these impacts are entirely driven by housing, where kickbacks are common. Nearly 40 percent of housing beneficiaries reported paying a kickback (column 3), with audits reducing incidence by 5.5 percentage points. Consequently, the value of reported kickbacks paid fell by ₹1022 ($12, column 4). The impact on leakage is four times larger at ₹4,661 ($57), reflecting the fact that respondents also reported receiving (insignificantly) more money from the housing program in the first place. While part of the difference in point estimates could be noise, some individuals may have been uncomfortable reporting kickbacks, instead simply reporting higher retained benefits. We therefore use the point estimates for kickbacks and leakage as lower and upper bounds on the effect of audits on unsanctioned payments to local implementers. Treatment effects are large in magnitude, representing a 17 percent reduction in both kickbacks and leakage.

In contrast, we find no impact on the likelihood an individual is a ghost beneficiary (column 2) – both for workfare (where ghosts are very common; 71 percent of control beneficiaries) and for housing (where they are rare; 13.5 percent of control beneficiaries).

To summarize, audits left housing registrations and materials benefits expenditures unchanged (which, combined with the null effect on the share of ghost beneficiaries implies no change in access to the program) while graft in the form of kickbacks reduced. Thus, from the government's perspective audits were a "win-win" proposition for this program, with monitoring transferring government resources from local implementers to intended recipients.

The picture for workfare is less promising. The reduction in administrative wage expenditure, combined with a null effect on ghost work's *share* of expenditure, implies absolute leakage in the form of ghost work declined, as did genuine payments to workers. Table 5 uses the voter roll sample to assess impacts on beneficiary access more directly, with Appendix Table A14 reporting treatment-on-the-treated effects. The first two columns study extensive margin effects on workfare participation during the follow up period. Workfare access in the general population was low, at 6 percent. Column 1 shows that treatment reduced self-reported workfare participation by 2.5 percentage points, or 41 percent of the control mean. Column 2 shows that these results are robust to expanding the definition of workfare access to include those who reported receiving program payments without working. Finally, column 3 shows that treatment reduced reported workfare wages received by ₹58, or 44 percent of the control mean. These results are larger than, but broadly in line with, the decline in workfare wage expenditure we observe in the survey sample administrative data (Table A15, column 2).

While reduced beneficiary access is problematic in light of workfare's stated aims, audits' total effect on leakage depends on how funds reallocated from wages to materials were used. It is possible that part of the reallocation reflected expenditure to suppliers made in a genuine effort to build higher quality assets (broadly consistent with program aims). Or, reallocation could reflect higher markups paid to contractors and materials suppliers to compensate for the fact that materials payments are often extremely delayed (inconsistent with program aims). A final possibility is that reallocated funds were captured by local implementers. Since implementers' revealed preference

was to direct more expenditure towards wages absent audits, an increase in funds capture would likely be offset by other costs, such as longer waiting times, increased hassle, or reduced electoral returns. We have limited ability to characterize how expenditure reallocated to materials was used. However, given workfare's prioritization of wage expenditure over asset generation and the fact that audits did not reduce the number of projects, audits appear to have had largely undesirable effects on workfare implementation.

5.4 Aggregate Impacts and Cost Effectiveness

We now use data from our experiment, which covered over a third of all audits implemented in FY 2019-20, to estimate statewide effects on program expenditure, leakage, and benefits received by citizens. To do this, we make the following choices: first, we focus on *citizen-facing* benefits (workfare wages, housing-linked labor benefits, and housing materials payments) setting aside audits' effects on workfare materials expenditure. Second, we assume that the treatment-on-the-treated effects on administrative data-reported expenditures in Table A11 (for labor benefits) and Table A9 (for housing materials benefits) apply to all audited panchayats. Third, we assume that rates of ghost work and ghost houses documented in Figure 1 apply statewide and were unaffected by audits. Finally, to account for kickbacks, we assume the treatment-on-the-treated effects in Panel B of Table A13 apply statewide, using the point estimate for kickbacks as a lower bound and the point estimate for leakage as an upper bound.[25] Table 6 reports results. Applying our point estimates to 2,645 audited panchayats and the 23 months in the post period implies audits reduced workfare wage expenditure by US$9.7 million, of which US$3.1 million was lost benefits for citizens and US$6.6 million was reduced ghost work. Point estimates for housing indicate US$9.7 million more sent to citizens at the cost of US$1.6 million more in ghost expenditure. We caveat, however, that these estimates are driven by insignificant positive point estimates on housing materials expenditure. Alongside, kickbacks under the housing program fell by $11.5 to $39.2 million. By comparison, the Social Audit Unit spent just $486,000 implementing social audits in FY 2019-20. Thus, the program's impacts on both workfare and housing far exceeded the cost of implementation.

6 Mechanisms

6.1 Reductions in Workfare Do Not Reflect Reduced Citizen Demand

There are two reasons why audits may have reduced access to workfare: reduced citizen demand or reduced effort by local implementers to provide workfare jobs. We included a detailed module on

[25] We use administrative data to calculate the total number of housing beneficiaries receiving payment per audited panchayat in the post period, and scale accordingly.

workfare demand in our follow up survey in order to differentiate between these two channels. Table 7 reports impacts on demand in the pooled sample (Panel A) and for the voter roll sample, workfare beneficiaries, and housing beneficiaries separately (Panel B). The first column reports impacts on the reservation wage.[26] Later on in the survey we offered respondents the opportunity to have our field team submit their (and/or family members') names to local officials, indicating that they were interested in participating in workfare at the going wage. To ensure respondents felt this question had real stakes, we collected the names and contact numbers of those interested and delivered them to officials at the block and panchayat level at the end of field activities. Column 2 reports impacts on the total number of household members the respondent wished to register for workfare. While the workfare reservation wage was unchanged (consistent with the fact that audits did not change the nature of workfare labor), treated individuals were interested in registering 13 percent more names. Panel B shows that the increase in workfare demand is largest among housing beneficiaries, who received better treatment under that program and asked to register 24 percent more names. This is consistent with the hypothesis that better experiences with local implementers increased demand for government services. Taken together, the results in column 2 of Table 7 refute the hypothesis that citizens demanded less workfare in the wake of audits. We therefore infer that expenditure changes primarily reflect reduced effort to provide workdays by local implementers.

6.2 Audits Disciplined Behavior of Contracted Implementers

Why did local implementers reduce graft in response to audits? A first possibility is that audits increased citizens' awareness of their rights and entitlements – a specific goal of the program. Better knowledge could improve citizens' ability to negotiate with officials or demand better services.[27] Column 3 of Table 7 tests this hypothesis by studying treatment effects on a GLS-weighted index of citizen awareness of workfare, with Appendix Table A16 reporting impacts on index components. While awareness and information was far from complete, audits had no impact on this family of outcomes.

A second possibility is collective action and bottom-up pressure. Audits made implementer misconduct salient both via the process of benefits verification, and via the public hearing (gram sabha) that took place at the end of an audit. These hearings may have informed citizens about misconduct and facilitated coordination among concerned community members. Indeed, one of the original aims of the program was to cultivate this type of collective action (Pande and Dubbudu, 2017). Appendix Table A17 shows that individuals in treatment panchayats were significantly more likely to remember that a gram sabha happened in their community, and report attending it themselves. This is especially true for workfare and housing beneficiaries, who had an explicit stake

[26] When eliciting this wage, a randomly selected half of respondents were given the example of working on a public asset, like a road or canal. The other half was randomly primed with an asset on their own land like a cattle shed. This randomization is orthogonal to the main audits treatment.

[27] This is not a given. In an earlier study in Bihar, Dutta et al. (2014) find that an information campaign with similar aims moved knowledge, but not workfare outcomes.

in the audit findings. This indicates scope for bottom-up pressure. We hypothesize that this channel will be especially salient for elected officials, who are directly accountable to citizens through the polls and faced an election in the last quarter of 2021, roughly 2 years after study audits.

A final possibility is top-down monitoring: if local implementers fear punishment from the state, they may change their behavior even absent citizen pressure. Beyond formal sanction, bureaucrats can punish local implementers employed by the state by exploiting career concerns, e.g. by opting not to re-appoint those on short-term contracts or offering undesirable transfers. By contrast, bureaucrats do not have easy means (nor incentives) to punish elected officials and social audits rarely trigger criminal prosecution.

To assess the relevance of top-down and bottom-up monitoring, we study audits' impacts on implementer behavior by type (elected versus contracted). We do this by focusing on kickbacks, since our survey recorded who in the community solicited them. We use these data to construct one dummy variable identifying respondents who reported giving a kickback to an elected official, and a separate dummy identifying respondents who reported kickbacks to contracted officials.[28] Table 8 shows that the reduction in kickbacks paid under the housing program is entirely driven by requests from contracted implementers.[29]

6.3 Elections and the Lasting Effects of Audits

Audits could impact electoral outcomes in several ways. A first channel relates to information disclosure: evidence from Brazil (Ferraz and Finan, 2008), Mexico (Larreguy et al., 2020), Puerto Rico (Bobonis et al., 2016) and South Africa (Berliner and Wehner, 2022) indicates that voters respond to "timely" pre-election audit results, punishing dishonest politicians and rewarding honest (or savvy) ones. Audits may have also reduced the perceived return to holding elected office, especially for politicians prone to graft. Even though elected officials did not change their own kickback-taking behavior under the housing program, they may have been less able to extract payments from contracted officials who did change their behavior.[30] Moreover, dishonest elected officials likely suffered losses associated with reduced ghost work under workfare. These channels would reduce dishonest politicians' expected returns to running for office, and possibly increase expected returns among honest politicians (who may anticipate a better chance of winning office). To test this hypothesis, we use electoral data to assess whether the 2016 incumbent local leader

[28] Elected officials include the panchayat leader, called the mukhiya or sarpanch, panchayat ward members, and panchayat committee members. Contracted officials include the panchayat sachiv, who is responsible for maintaining community records, the panchayat rozgar sewak (responsible for administering workfare) the aawas sahayak (responsible for housing verification), mates (who help run workfare work sites), and officials at the block office. Some responses – such as "broker" or "people in the village" cannot be classified.

[29] The most common culprit in this category is the aawas sahayak, who is tasked with ensuring that PMAY-G houses are constructed per program rules, including uploading photographic proof of home construction to the program website. At the time of the study, aawas sahayaks held three-year contracts.

[30] Qualitative evidence indicates that contracted officials sometimes collect kickbacks at the behest of, or pay a portion of kickback revenue to, elected officials.

(*mukhiya*) (a) ran and (b) won. These results reflect the net effect of informing voters *and* reducing future returns to graft. We also study impacts on the number of challengers, which shuts down the voter information channel.[31]

Table 9 reports results. Audits reduced the likelihood of incumbents running by 5 percentage points (6 percent) and the number of challengers by 0.37 (5 percent). We find no impact on the likelihood the incumbent won re-election, though point estimates are non-trivial relative to the dependent variable mean (a 2 percentage point reduction against an 18 percent re-election rate in the control group). While we cannot directly measure entry/exit of more versus less honest candidates, we find two pieces of evidence that suggest audits induced favorable electoral selection. First, Appendix Table 10 uses results from audits that took place after the RCT (which, per Appendix Table A6, are uncorrelated with treatment) to assess the FY 2019-20 audits' longer-term impacts on program implementation. We use the number and monetary amount of issues as proxies of graft – importantly, in the control group these outcomes are significantly correlated with graft measured in our follow-up survey (Appendix Figure A5). We find that treatment significantly reduced the number (but not value) of issues reported in non-experimental audits post FY 2019-20. Notably, these reductions are concentrated in post-election years (FY 2022-23 and FY 2023-24). Consistent with this, we find that citizens in treatment panchayats were more satisfied with their leaders after elections – as measured by reports that local leaders could help them receive government benefits in our (post-election) follow-up survey (Table 9, column 5). Further, recall that citizens in treated panchayats were more interested in participating in workfare, despite no change in the reservation wage (Table 7), bolstering the hypothesis that citizens were overall more satisfied with local service delivery.

6.4 Understanding Differing Impacts on Workfare and Housing

Taken together, we find that audits significantly affected the actions of both contracted and elected officials, with very different implications for India's workfare and housing programs. Why did implementers reduce effort to deliver services under workfare, but not housing? While the two programs differ in many ways, considering programmatic differences in how implementers are monitored and the amount of discretion they are afforded is instructive. The housing program is much simpler, meaning implementers have less discretion and are easier to monitor. Beneficiary lists are determined through large scale surveys and verified by community meetings. State- and district-level officials use these lists to assign annual targets (down to specific beneficiary names) to panchayat-level implementers. Once assigned, progress against these targets are regularly monitored through administrative data. Graft occurs precisely where implementers do have discretion, monitoring is

[31]This assumes potential challengers in audited areas perceived a greater risk of future audits. This is consistent with evidence from Brazil (Avis et al., 2018) and qualitative feedback from implementers. If instead treatment left the perceived risk of future audits unchanged, then effects on challengers' behavior would only reflect changes in the perceived likelihood of winning. Our analysis of electoral effects was not pre-specified.

limited, and there is a natural holdup problem: it is left to implementers to decide whether to certify a beneficiary's progress building their home, which is required to open up access to the second and third installments under the program. Audits significantly improved the government's ability to monitor at this stage – it is therefore intuitive that kickbacks declined. We further hypothesize that total expenditures remained unchanged because the state was already robustly monitoring progress against spending targets.

Under workfare, annual plans at the beginning of the fiscal year fix panchayat budgets. Workfare implementers are monitored on progress against these plans; the government also regularly monitors whether public works projects are actually built, but has limited ability to verify asset quality. Given the program's demand-driven ethos implementers have significant discretion over *who* in the community receives labor benefits – which helps enable ghost work. Panchayat-level officials also had non-trivial leeway to shift expenditure between the labor and material budgets during our study period. This is both because the labor expenditure share is only tracked at the district level, and because the 60 percent labor expenditure floor was less binding during our study period. This was due to the large volume of housing-linked labor payments (which, unlike workfare expenses, are tied to no MGNREGS materials expenditure). Audits made it much easier for the state to monitor ghost work, but did not do much to improve monitoring of the materials budget (since verifying asset quality and materials needed to build an asset requires significant expertise). Our results are consistent with the hypothesis that audits reduced the (graft-related) relative return to generating wage expenditure, with implications for both leakage and citizen access. One question here is why implementers adjusted on the margin of total labor expenditure, leaving the average rate of ghost work unchanged. Insofar as providing genuine work is costly (due to the need to publicize workers, process paperwork, and supervise them on site), officials may have scaled back their effort in light of lower overall returns to generating labor expenditures.

7 Conclusion

We find evidence that women-led social audits of two of India's largest social protection programs shift critical aspects of government service delivery. First, graft in the rural housing program declined significantly, driven by lower incidence and value of kickbacks paid to local officials. These changes are important in magnitude, both at the individual level – assignment to treatment reduced kickbacks paid by housing program beneficiaries by ₹1022 (US$12), and at a macro level – our experimental estimates imply that audits reduced statewide graft in the housing program by US$11.5–39.2 million, relative to total audit implementation costs of US$486,000.

Results for the workfare program are less promising. Total expenditure on wage payments declined, offset by a (statistically insignificant, but equal in magnitude) increase in materials expenditure. As a result, the labor share of total expenditure under the program fell. This reflects both reduced benefits retained by citizens (US$3 million, in our statewide calculations) and reduced graft

(US$7 million). However, this improvement may have been offset by increased graft in materials expenditure (where misallocation is more effortful, but more difficult to detect).

Reduced access to workfare could reflect either less citizen demand or less effort to provide work on the part of local implementers. Evidence indicates reductions reflect reduced implementation effort. Importantly, citizens' demand for workfare increased, with large increases in demand evident among housing beneficiaries, who experienced less expropriation under audits.

Finally, we find evidence that audits have prolonged effects by impacting electoral outcomes. Audits reduced the number of candidates (including non-incumbents) vying for local elected office, in line with the hypothesis that reduced opportunities to engage in graft dissuaded dishonest contenders. Consistent with this hypothesis, our follow-up survey – conducted after elections had been decided and new leaders were in power – finds that citizens in audited communities were significantly more likely to report that officials were capable of helping them access government benefits. Moreover, treatment significantly reduced the number of issues recorded in future, post-election rounds of social audits, suggesting better behavior once new leaders were in power.

Our results provide some of the first evidence that monitoring raises a tradeoff between deterring graft and supporting program reach through effort of local implementers. Pointing to the importance of institutional features, this tradeoff is only present for workfare, where implementers have significantly more discretion over program budget and beneficiary selection. Alongside, we highlight the promises and pitfalls of the largest citizen-led auditing program in the world, which touches the lives of nearly a billion citizens every year. One limitation of our research is that it reflects "first effects" of audits. Deterrence power could reduce over time, if, for example, implementers learn that there are few consequences for bad behavior. On the other hand, our results could understate the positive effects of audits insofar as audits induce favorable selection of elected leaders. Understanding the longer-term dynamics of such programs is an important area for future research.

References

Acemoglu, D. and T. Verdier (2000). The choice between market failures and corruption. *American Economic Review 91*(1), 194–211.

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.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Avis, E., C. Ferraz, and F. Finan (2018). Do government audits reduce corruption? estimating the impacts of exposing corrupt politicians. *Journal of Political Economy 126*(5), 1912–1964.

Banerjee, A., R. Hanna, and S. Mullainathan (2012). 27. corruption. In *The Handbook of Organizational Economics*, pp. 1109–1147. Princeton University Press.

Berliner, D. and J. Wehner (2022). Audits for accountability: evidence from municipal by-elections in south africa. *The Journal of Politics 84*(3), 1581–1594.

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

Cuneo, M., J. Leder-Luis, and S. Vannutelli (2023). Government audits. Technical report, National Bureau of Economic Research.

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.

Ferraz, C. and F. Finan (2008). Exposing corrupt politicians: the effects of brazil's publicly released audits on electoral outcomes. *The Quarterly journal of economics 123*(2), 703–745.

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

Fischer, H. W. and S. S. Ali (2019). Reshaping the public domain: Decentralization, the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), and trajectories of local democracy in rural India. *World Development 120*, 147–158.

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

International Labour Office (2024). *World Social Protection Report 2024–2026: Universal Social Protection for Climate Action and a Just Transition*. Geneva: International Labour Office.

J-PAL (2019). Governance review paper j-pal governance initiative.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Larreguy, H., J. Marshall, and J. M. Snyder Jr (2020). Publicising malfeasance: when the local media structure facilitates electoral accountability in mexico. *The Economic Journal 130*(631), 2291–2327.

Leff, N. H. (1964). Economic development through bureaucratic corruption. *American behavioral scientist 8*(3), 8–14.

Leys, C. (1965). What is the problem about corruption? *The Journal of Modern African Studies 3*(2), 215–230.

Mathew, S. and M. Moore (2011). State incapacity by design: Understanding the bihar story. *IDS Working Papers 2011*(366), 1–31.

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.

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

Pande, S. and R. R. Dubbudu (2017). Citizen oversight and india's right to work program: What do the social auditors say. *Accountability Research Center, Accountability Working Paper 1*.

PLFS (2024). Annual Report: Periodic Labour Force Survey (PLFS), July 2023 - June 2024. Annual report, Ministry of Statistics and Programme Implementation, National Sample Survey Office. Data period: July 2023 - June 2024.

Ravallion, M. (2019). Is a decentralized right-to-work policy feasible? Technical report, National Bureau of Economic Research.

Wong, W. N. (2024). Optimal monitoring and bureaucrat adjustments. Working paper, University of Chicago, Harris School of Public Policy.

Zamboni, Y. and S. Litschig (2018). Audit risk and rent extraction: Evidence from a randomized evaluation in brazil. *Journal of Development Economics 134*, 133–149.

Tables and Figures

Table 1: Audit Implementation: Results from Process Monitoring Surveys

	Mean	N
Panel A: Auditor Characteristics, Training, and Needs		
Female	0.80	46
Age	31.39	46
Education: Less Than Class 12	0.37	46
Education: Class 12	0.39	46
Education: Any Postsecondary	0.24	46
Travel Time: Home to Audited Community (Minutes)	174.24	46
Received Training	1.00	46
Feels Unsafe Conducting Audits	0.04	46
Feels Somewhat Safe Conducting Audits	0.59	46
Feels Safe Conducting Audits	0.37	46
Reported Need for More Training	0.61	46
Panel B: Audited Beneficiaries Survey		
Participated in Workfare During Audit Period	0.46	360
Ever Demanded Workfare from Implementers	0.14	362
Would Like Workfare Work	0.61	357
Got Paid Without Working	0.25	355
Aware of Audit Team	0.62	362
Visited by Audit Team	0.46	359
Aware of Community Meeting (Surveys One Day Prior)	0.30	185
Attended Community Meeting (Surveys One Day Post)	0.15	176
Panel C: Auditor Actions During Shadow Visits		
Verified Whether the Respondent Worked	0.94	380
Asked About the Number of Days Worked	0.27	300
Asked Respondent if Wages Were Paid to Bank Account	0.44	300
Asked Whether the Household Had a Job Card	0.78	380
Recorded Workfare Demand	0.06	304
Informed Respondents About Community Meting	0.37	380
Panel D: Community Meeting Observations		
Community Meeting Happened	1.00	13
Attendance: 1-50	0.15	13
Attendance: 51-100	0.46	13
Mukhiya (Elected Local Leader) Present	0.92	13
PRS (Contracted Workfare Implementer) Present	0.92	13
District Official Present	0.54	13
Time Discussing Workfare: Greater Than 20 Minutes	0.46	13
Time Discussing Housing: Greater Than 20 Minutes	0.38	13
Number Citizens Publicly Stating Workfare Grievances	3.85	13
Number Citizens Publicly Stating Housing Grievances	2.38	13
Enumerator Felt Mukhiya Had Undue Influence on Proceedings	0.38	13

Notes: All data come from process monitoring visits conducted in 24 panchayats in 13 districts. Data for Panel A come from a survey of a randomly-selected subset of beneficiaries selected for workfare verification as part of social aduits. We aimed to interview half these respondents one day before the audit community meeting and the other half one day after the community meeting. Data for Panel B come from shadow visits in which trained enumerators accompanied social auditors on workfare verification visits. Data for Panel C come from community meeting observations collected in 1 panchayat per district (the team completed fieldwork pre-meeting in the other sampled panchayats. Data for Panel D come from surveys of social auditors in process monitoring panchayats. We limit attention to frontline auditors, called village resource persons, excluding their supervisors, called master resource persons. The mukhiya is the elected leader of the panchayat. The PRS is a contracted official responsible for workfare administration.

Table 2: Impact of Treatment on Administrative Expenditure on Citizen-Facing Benefits Under Workfare and Housing

	Diff-in-Diff	ANCOVA	ANCOVA (M)	Diff-in-Diff	ANCOVA	ANCOVA (M)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Administrative Data - Workfare and Housing Labor Benefits						
	Houshold Demand Registered			Wage Expenditure		
Treatment	-0.42			-1.41		
	(3.20)			(7.61)		
Treatment × Anticipation Period	3.74	3.67	5.41	14.30	14.03	18.56
	(4.10)	(4.16)	(4.19)	(13.86)	(14.25)	(14.51)
Treatment × Post Period	-4.85**	-4.91**	-3.15	-14.52**	-14.76**	-10.18*
	(2.34)	(2.32)	(2.21)	(6.17)	(6.07)	(5.20)
Control Mean: Anticipation Period	116.20	116.20	116.20	293.05	293.05	293.05
Control Mean: Post Period	133.94	133.94	133.94	318.03	318.03	318.03
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512
Panel B: Administrative Data - Housing Registration and Materials Benefits						
	Housing Registrations			Housing Materials Expenditure		
Treatment	-23.25			-25.59**		
	(14.65)			(12.34)		
Treatment × Anticipation Period	23.53	24.58	32.73*	-3.68	2.19	1.93
	(19.09)	(19.34)	(19.42)	(15.27)	(15.81)	(15.21)
Treatment × Post Period	-0.65	0.65	8.86	6.16	12.32	12.08
	(25.81)	(25.72)	(23.63)	(20.79)	(20.11)	(19.70)
Control Mean: Anticipation Period	38.16	38.16	38.16	331.34	331.34	331.34
Control Mean: Post Period	1096.51	1096.51	1096.51	957.19	957.19	957.19
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512

Notes: Column headers list outcome variables of regressions. The first and fourth columns run a differences-in-differences specification, including data from April 2018 until April 2021. The remaining columns use data from April 2019 until April 2021 and control for the pre-period (April 2018-March 2019) outcome mean. ANCOVA (M) specifications control for household labor demand registered, total wage expenditure, housing registriations, and housing materials expenditure. All expenditures are measured in ₹1,000s. The sample is restricted to blocks where at least one audit happened. All regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table 3: Assessing Consequences of Declining Wage Expenditure Under Workfare

	A. Decomposing Wage Expenditure			*B. Assessing Reallocation to Materials Expenditure*		
	Total Wage Expenditure	Workfare Wage Expenditure (₹1,000s)	Housing Wage Expenditure (₹1,000s)	Workfare Materials Expenditure (₹1,000s)	Total Wage Expenditure Share	Workfare Wage Expenditure Share
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment × Anticipation Period	17.510	16.786	0.725	9.247	-0.006	-0.000
	(11.495)	(11.344)	(2.937)	(10.560)	(0.010)	(0.017)
Treatment × Post Period	-9.146*	-7.417*	-1.728	13.572	-0.013**	-0.030***
	(5.089)	(4.362)	(2.058)	(10.133)	(0.006)	(0.009)
Control Mean: Anticipation Period	232.45	157.02	75.43	97.87	0.84	0.68
Control Mean: Post Period	293.10	155.66	137.44	181.80	0.74	0.57
Number of Panchayats	2101	2101	2101	2101	2078	2064
N	52512	52512	52512	52512	6199	5694

Notes: Column headers list outcome variables of regressions. The time period spans from April 2019 until April 2021. Material and wage expenditure is available at the panchayat × month level, but we calculate wage expenditure shares at the panchayat × time period level (pre, anticipation, implementation, and post-implementation period) to avoid missing values and smooth outliers driven by sporadic invoicing of materials expenditure. The sample is restricted to blocks where at least one audit happened. All regressions control for pre-period mean of the outcome variable (when available), alongside the mean pre-implementation housing material expenditure sanctioned, the mean pre-implementation housing material expenditure, the mean pre-implementation household demand registered, and the mean pre-implementation total wage expenditure (including workfare and housing). The total wage expenditure share divides total wage expenditure (for both workfare and housing) by the sum of total wage and workfare materials expenditure. The workfare wage expenditure share divides workfare wage expenditure by the sum of workfare wage ependiture and workfare materials expenditure. There were two panchayats that had a missing value for wage expenditure shares due to no expenditures in the pre period. We recode these values to the pre-period sample mean. Additionally, all regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table 4: Impact of treatment on program graft and misallocation

		Index components				
	Misallocation index	Any Ghost Payment	Paid Kickback	Kickback	Payment Kept by Respondent (in Index Reversed)	Leakage
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Overall Effect						
Treatment	-0.087**	-0.012	-0.014	-464.287**	1493.282	-2143.026**
	(0.037)	(0.017)	(0.014)	(229.769)	(1143.247)	(1063.158)
Control Mean	-0.00	0.44	0.24	2935.68	43637.91	14588.86
N	4341	4341	4193	4155	3999	3884
Panel B. By Beneficiary Type						
Treatment × Housing	-0.197***	-0.011	-0.055**	-1021.602**	3481.314	-4661.192**
	(0.068)	(0.018)	(0.027)	(503.023)	(2420.060)	(2345.158)
Treatment × Workfare	0.006	-0.013	0.020	10.151	-291.498	-19.760
	(0.037)	(0.027)	(0.013)	(119.547)	(548.388)	(661.628)
Housing	-0.151***	-0.575***	0.368***	5897.321***	87555.593***	22083.530***
	(0.057)	(0.025)	(0.020)	(396.864)	(1884.101)	(1914.243)
P-value: Treat × Workfare=Treat × Housing	0.01	0.95	0.02	0.06	0.14	0.07
Control Mean (Workfare)	0.07	0.71	0.06	95.40	752.51	4556.16
N	4341	4341	4193	4155	3999	3884

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. All columns include outcomes from stacked data, where there is one observation per beneficiary for each program type as relevant. Panel A shows the overall effect. Panel B shows the decomposed effect by program type. The *misallocation index* is a GLS standardized index, which includes *any ghost payment*, *paid kickback*, *payment kept by respondent*, and *leakage* as components, with all components coded so higher values correspond to greater misallocation. Kickbacks, payment kept by respondent, and leakage are all measured in ₹s and topcoded at the 99th percentile. All regressions include block fixed effects and control for gender and a dummy to identify individuals in the housing sample. Details of variable construction are described in Appendix B. Clustered standard errors at the panchayat level. ***$p\leq0.01$, **$p\leq0.05$, *$p\leq0.1$.

Table 5: Impact of treatment on citizens' access to workfare

	Participated in Workfare	Participated in or Paid for Workfare	Total Workfare Wages: Apr. 2019 - Mar. 2020
	(1)	(2)	(3)
Treatment	-0.025**	-0.026**	-58.419*
	(0.011)	(0.012)	(31.516)
Control Mean	0.061	0.073	131.687
N	1868	1868	1835

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. The sample is restricted to the voter roll sample. All regressions includes block fixed effects and control for gender. Total workfare wages is reported in ₹s and top-coded at the 99th percentile. Clustered standard errors at the panchayat level. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table 6: Statewide Impacts of Audits on Expenditure and Misallocation

	Change in Government Expenditure:				
	Total	Sent to Citizens	Ghost Expenditure	Change in Kickbacks	Total Change in Leakage (Ghosts + Kickbacks)
	(1)	(2)	(3)	(4)	(5)
Workfare	-9,729,610	-3,113,475	-6,616,135	0	-6,616,135
Housing	11,225,766	9,654,158	1,571,6070	-11,495,218 to -39,187,243	-9,923,611 to -37,615,636
Total	1,496,155	6,540,683	-5,044,528	-11,495,218 to -39,187,243	-16,539,746 to -44,231,771

Notes: All figures are in US$, assuming an exchange rate of ₹82 per US$. We assume audits had no impact on rates of ghost houses, ghost work and workfare materials expenditure. We assume rates of ghost houses and ghost work reported in Figure 1 apply statewide and assume that the treatment-on-the-treated effects on administrative expenditures in Table A11 (for workfare and housing wage payments) and Table A9 (housing materials payments) apply to all audited panchayats. Estimates for the change in kickbacks under housing use point estimates from Table A13, with impacts on reported kickbacks as a lower bound and impacts on leakage as an upper bound. To create statewide estimates, for administrative expenditure (columns 1-3) we take monthly treatment-on-the-treated estimates and multiply them by 23 (the number of months in the post period) and 2,645 (the number of audited panchayats in FY 2019-20). For kickbacks (column 4) we take treatment on the treated effects and multiply them by 243 (the number of housing beneficiaries receiving payment per audited GP in the post period) and then 2,645. The total budget of Bihar's social audit society in FY 2019-20 was $486,000.

Table 7: Impact of treatment on demand for and Awareness of Workfare

	Workfare Reservation Wage	Workfare Demand	Workfare Awareness Index
	(1)	(2)	(3)
Panel A. Overall Effect			
Treatment	0.134	0.060**	0.029
	(4.799)	(0.027)	(0.030)
Control Mean	315.46	0.48	-0.00
N	6170	6232	6278
Panel B. By Beneficiary Type			
Treatment × Housing	-6.607	0.104***	0.037
	(8.093)	(0.036)	(0.051)
Treatment × Workfare	1.766	0.070	0.020
	(7.102)	(0.048)	(0.037)
Treatment × Voter Roll	5.315	0.001	0.032
	(7.698)	(0.042)	(0.053)
Housing	21.731***	-0.120***	-0.217***
	(7.665)	(0.042)	(0.038)
Voter Roll	34.747***	-0.077*	-0.275***
	(7.115)	(0.042)	(0.044)
Control Mean (Workfare)	295.91	0.55	0.19
N	6170	6232	6278

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. All columns include outcomes from stacked data, where there is only observation per program for each beneficiary type as relevant. Panel A shows the overall effect. Meanwhile, Panel B shows the decomposed effect by beneficiary type. All regressions include block fixed effects and control for gender, individuals drawn from housing sample and invididuals drawn from the voter roll sample. The reservation wage is the minimum wage at which the survey respondent reports being willing to work for workfare, top-coded to ₹550 for those who state they are never willing to perform such work. Workfare demand is the number of household members whose names (including him/herself) the respondent requests being given to implementers to receive workfare. Workfare awareness index is a GLS standardized index, that includes dummy variables for awareness of workfare program, aware of a document called job card, aware of the number of days entitled to work, aware of payment method, aware of daily wage, and aware of eligibility to work as components. Clustered standard errors at the panchayat level. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table 8: Impact of Treatment on Kickbacks by Type

	Paid Kickback to Elected Official (1)	Paid Kickback to Contracted Official (2)
Panel A. Overall Effect		
Treatment	-0.003	-0.014
	(0.012)	(0.009)
Control Mean	0.13	0.10
N	4190	4190
Panel B. By Beneficiary Type		
Treatment × Housing	-0.003	-0.035*
	(0.025)	(0.020)
Treatment × Workfare	-0.003	0.004
	(0.010)	(0.006)
Housing	0.198***	0.193***
	(0.019)	(0.015)
P-value: Treat × Workfare=Treat × Housing	0.98	0.06
Control Mean (Workfare)	0.03	0.01
N	4190	4190

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. All columns include outcomes from stacked data, where there is one observation per program for each beneficiary type as relevant. The sample is restricted to beneficiaries of workfare and housing. Panel A shows the overall effect. Meanwhile, Panel B shows the decomposed effect by beneficiary type. All regressions include block fixed effects and control for gender and individuals drawn from the housing sample. Details of variable construction are described in Appendix B. Clustered standard errors at the panchayat level. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table 9: Impact of Treatment on Electoral Outcomes and Citizen Satisfaction with Local Leaders

	Electoral Data				Endline Survey
	Incumbent Ran	Number Challengers	Incumbent Won	Incumbent Vote Share	Thinks Local Leaders Can Help in Receiving Benefits
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.045*	-0.369**	-0.017	-0.004	0.032**
	(0.025)	(0.180)	(0.023)	(0.008)	(0.016)
Control Mean	0.75	7.68	0.18	0.17	0.45
	1792	1788	1804	1788	6075

Notes: Regressions in columns 1-4 are the panchayat level, while the regression in column 5 is at the individual level. The endline survey regression includes additional controls for respondent gender and whether the individual is drawn from the voter or workfare sample. All regressions include block fixed effects. The electoral outcomes refer to 2021 panchayat elections. Standard errors are in parentheses, clustered at the block level in columns 1-4 and the panchayat level in column 5. ***$p<0.01$, **$p<0.05$, *$p<0.1$.

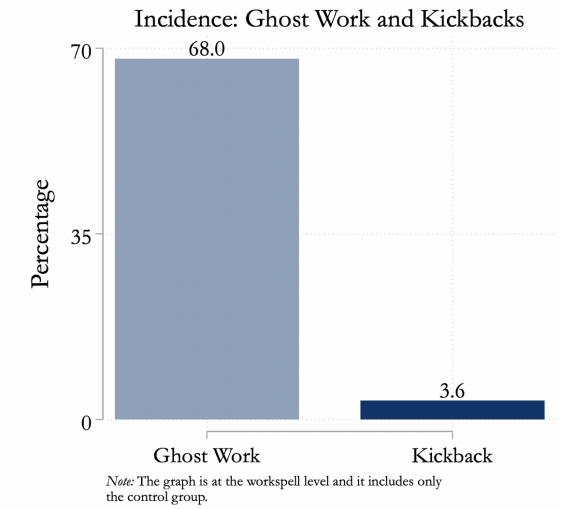
Table 10: Impact of Treatment on Number and Value of Issues in Future Audits

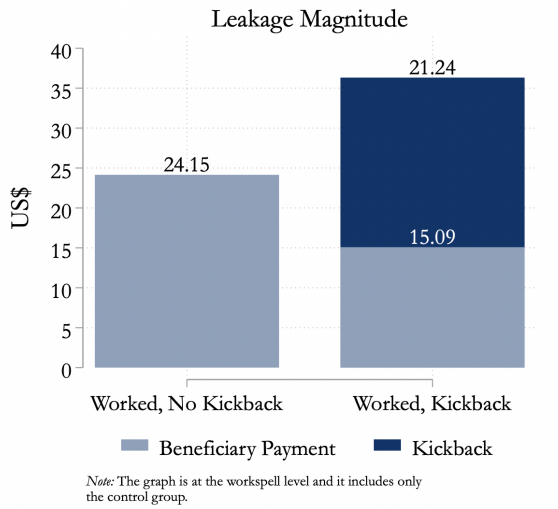
	Number of Issues Post FY19-20			Value of Issues Post FY19-20		
	Total Post FY19-20	Pre Panchayat Elections	Post Panchayat Elections	Total Post FY19-20	Pre Panchayat Elections	Post Panchayat Elections
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Full RCT Sample (Blocks Where Audits Happened)						
Treatment	-0.12*	0.28	-0.16**	489.02	5655.67	90.77
	(0.07)	(0.19)	(0.07)	(372.65)	(5472.76)	(72.12)
Control Mean	6.15	7.49	6.05	565.24	679.50	545.06
N	1998	509	1995	1998	509	1995
Panel B. Endline Survey Sample						
Treatment	-0.32**	-0.12	-0.33**	-34.56	-581.08	7.55
	(0.15)	(0.15)	(0.15)	(146.90)	(439.83)	(144.30)
Control Mean	6.51	8.76	6.36	514.53	702.83	468.24
N	273	59	273	273	59	273

Notes: Regressions are at the panchayat level and include block fixed effects. The value of issues is measured in ₹1,000s. Panchayat elections concluded in December 2021. Thus the pre-panchayat election period includes FY21-22, while the post election period includes FY22-23 and FY23-24. Heteroskedasticity robust standard errors in parentheses. ***$p<0.01$, **$p<0.05$, *$p<0.1$.

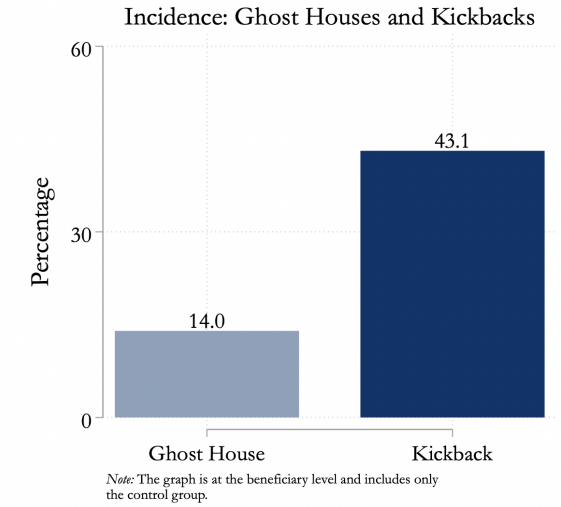
Figure 1: Status Quo Leakage in Workfare and Housing

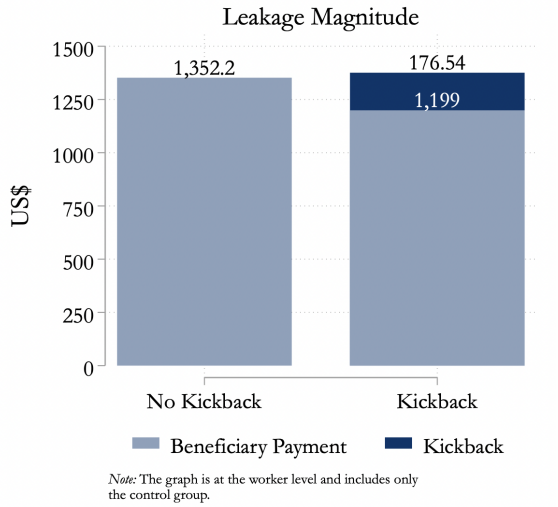
A. Workfare Work Spells





B. Housing Beneficiaries





Notes: Data come from the endline survey control group. Panel A includes workfare work engagements. Incidence reports the share of all work engagements that are "ghost work" (unconfirmed by the listed beneficiary) and the share of work engagements where beneficiaries reported paying a kickback. Leakage magnitude limits the sample to work engagements where the beneficiary worked and graphs the total payment retained by the beneficiary (and, when relevant, the kickback paid to local leaders). Panel B includes all beneficiaries listed as having received labor payments for construction of a PMAY-G house. Incidence documents the share of respondents classified as "ghost houses" (where beneficiary states they do not have a PMAY-G home) and the share of respondents who reported paying a kickback; leakage magnitude limits the sample to beneficiaries who confirmed receipt of a PMAY-G house and graphs total payments received and kickbacks paid for the home.

A Appendix Tables and Figures

Table A1: Pre-Intervention Differences in Panchayat-Level Outcomes

	RCT Blocks Where At Least One Audit Happened		
	Control Mean	Treatment-Control Difference (Regression Adjusted)	Number of Panchayats
	(1)	(2)	(3)
Panel A: Administrative Data			
Person-days Demand Registered	96.563	-0.853	2101
	[128.399]	(3.496)	
Any Person-days of Work Provided	0.782	0.003	2101
	[0.413]	(0.008)	
Wage Expenditures (Workfare + Housing, ₹1,000s)	188.649	-1.415	2101
	[309.234]	(7.608)	
Workfare Material Expenditure (₹1,000s)	107.885	1.419	2101
	[327.734]	(6.653)	
Housing Material Beneifts Registered (₹1,000s)	492.296	-23.251	2101
	[1205.176]	(14.651)	
Housing Material Benefits Expenditure (₹1,000s)	506.367	-25.587**	2101
	[611.785]	(12.340)	
Panel B: 2011 Socio-Economic and Caste Census Data			
Total Population (1,000s)	13.336	0.226	1801
	[6.057]	(0.292)	
Percent Female	47.131	-0.018	1801
	[0.858]	(0.043)	
Percent With More Than Primary Education	17.715	0.148	1801
	[6.461]	(0.249)	
Percent Households with Highest Earner < ₹5,000	71.188	0.081	1801
	[11.740]	(0.482)	
Percent Households Economically Marginalized	62.073	-0.192	1801
	[11.276]	(0.470)	
Percent Households Relying on Manual Casual Labor	69.904	-0.175	1801
	[11.546]	(0.447)	

Notes: Sample limited to blocks where at least one audit happened. The second column shows regression adjusted differences between treatment and control panchayats. Panel A focuses on adminstrative data outcomes from the MGNREGS and PMAY-G web portals. Measures of registered demand and wage expenditure provided pool together workfare- and housing-linked benefits. All outcomes in Panel A are at the panchayat × month level. Panel B is at the panchayat level and reports community socio-economic characteristics measured in the 2011 Socio-Economic and Caste Census (SECC). Economically marginalized is defined as those households which have at least one SECC deprivation or are classified as poor. The number of observations in Panel B is smaller than our total RCT sample because we were not able to merge all our study panchayats to the SECC due to name changes and panchayat subdivision over time. The sample in Panel A is restricted to the pre-implementation period spanning between April 2018 and March 2019. All regressions include block fixed effects. Standard deviations are in brackets, robust standard errors clustered at the block level are in parentheses. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table A2: Pre-Audit Differences Across Panchayats – Blocks Where No Audits Happened and Non-RCT Blocks

	RCT Blocks Where No Audits Happened			Non-RCT Blocks		
	Control Mean	Treatment-Control Difference (Regression Adjusted)	Number of Panchayats	Non-Audited Mean	Audited-Non-Audited Difference (Regression Adjusted)	Number of Panchayats
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Administrative Data						
Person-days Demand Registered	100.157	-8.207**	1336	111.364	23.400***	5093
	[138.434]	(4.055)		[145.208]	(7.271)	
Any Person-days of Work Provided	0.818	-0.013	1336	0.853	0.039***	5093
	[0.386]	(0.011)		[0.354]	(0.011)	
Wage Expenditure (Workfare + Housing, ₹1,000s)	188.503	-13.030	1336	213.185	40.791***	5093
	[300.113]	(8.846)		[324.209]	(14.360)	
Workfare Material Expenditure (₹1,000s)	97.381	1.753	1336			
	[261.131]	(8.016)				
Panel B: 2011 Socio-Economic and Caste Census Data						
Total Population (1,000s)	12.424	0.452*	1151	12.881	-0.418*	4741
	[4.787]	(0.264)		[5.155]	(0.221)	
Percent Female	47.190	0.040	1151	47.376	0.032	4741
	[0.979]	(0.058)		[0.924]	(0.066)	
Percent With More Than Primary Education	19.403	-0.405	1151	17.025	1.208*	4741
	[6.850]	(0.277)		[7.747]	(0.663)	
Percent Households with Highest Earner < ₹5,000	67.971	-0.184	1151	71.329	-0.778	4741
	[13.459]	(0.572)		[12.373]	(0.885)	
Percent Households Economically Marginalized	58.525	0.431	1151	59.996	-0.539	4741
	[14.285]	(0.542)		[13.033]	(1.091)	
Percent Households Relying on Manual Casual Labor	71.791	-0.039	1151	70.345	-1.157	4741
	[10.528]	(0.558)		[11.072]	(0.823)	

Notes: The second column shows regression adjusted differences between treatment and control panchayats in RCT blocks where no audit happened. The fifth column shows regression adjusted differences between non-audited and audited panchayats outside the RCT sample. Panel A repports adminstrative data outcomes from the MGNREGS web portal. Measures of registered demand and wage expenditure provided pool together workfare and housing-linked benefits. All data in Panel A are at the panchayat × month level. Data on workfare material expenditure is only available for the RCT sample. Panel B is at the panchayat level and reports community socio-economic characteristics measured in the 2011 Socio-Economic and Caste Census (SECC). Economically marginalized is defined as those households which have at least one SECC deprivation or are classified as poor. *Material expenditure billed* and *Total poulation* are in thousands. The number of observations in Panel B is smaller than our total RCT sample because we were not able to merge all our study panchayats to the SECC due to name changes and panchayat subdivision over time. The sample in Panel A is restricted to the pre-implementation period spanning between April 2018 and March 2019. All regressions include block fixed effects. Standard deviations are in brackets, robust standard errors clustered at the block level are in parentheses. ***$p\leq0.01$, **$p\leq0.05$, *$p\leq0.1$.

Table A3: Tracking rate across treatment groups

	Dep. Variable: Took Survey	
	Beneficiary Sample	Voter Roll Sample
	(1)	(2)
Treatment	-0.001	-0.002
	(0.013)	(0.018)
Control Mean	0.78	0.58
N	5479	3288

Notes: Column headers list sample restriction. Each column is a separate regression of the outcome variable on a treatment dummy. The data used to run these regressions contains all final survey attempts. All regressions include block fixed effects. Cluster standard errors at the panchayat level in paranthesis. ***$p<0.01$, **$p<0.05$, *$p<0.1$

Table A4: Summary Statistics and Balance – Voter Roll Sample

	Control Mean	Regression Adjusted Difference	N
	(1)	(2)	(3)
Female	0.58	-0.02	1844
	[0.49]	(0.01)	
Age	42.50	-0.61	1844
	[15.35]	(0.70)	
General Caste	0.14	0.00	1844
	[0.34]	(0.02)	
Scheduled Caste or Scheduled Tribe	0.19	0.00	1844
	[0.39]	(0.02)	
Other Backward Clasess	0.30	-0.04	1844
	[0.46]	(0.03)	
Economically Backward Classes	0.37	0.04	1844
	[0.48]	(0.03)	
Hindu	0.91	-0.02	1844
	[0.29]	(0.02)	
Muslim	0.09	0.02	1844
	[0.29]	(0.02)	
Married	0.88	-0.00	1844
	[0.32]	(0.01)	
Widowed	0.07	-0.00	1844
	[0.26]	(0.01)	
No education	0.46	-0.01	1844
	[0.50]	(0.02)	
Primary education (up to class 8)	0.31	-0.02	1844
	[0.46]	(0.02)	
Secondary or higher education	0.23	0.04*	1844
	[0.42]	(0.02)	
Government connection	0.06	-0.00	1844
	[0.24]	(0.01)	
Household members (includes the respondent)	6.36	-0.16	1844
	[3.29]	(0.16)	
Lives in the location for more than 6 months a year	0.87	-0.01	1844
	[0.33]	(0.01)	
Number of assets owned (out of 12)	5.07	-0.00	1844
	[1.70]	(0.08)	
Working	0.56	-0.01	1844
	[0.50]	(0.02)	
Self-employed in agriculture	0.19	-0.00	1844
	[0.39]	(0.02)	

Notes: The first column shows the means of the outcome variable in the control group. The second column shows the regression coefficient of the outcome variable on an indicator variable for being treated, including block fixed effects. The sample is restricted to the voter roll sample. *Government connection* measures social connections, it equals one if the respondent or a member of their family holds a panchayat/block position, and zero otherwise. Clustered standard errors at the panchayat level in parentheses. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table A5: Summary Statistics and Balance – Beneficiary Sample

	Control Mean	Regression Adjusted Difference	N
	(1)	(2)	(3)
Female	0.54	-0.01	4195
	[0.50]	(0.01)	
Age	45.17	-1.16**	4195
	[14.59]	(0.51)	
General Caste	0.11	-0.01	4195
	[0.31]	(0.02)	
Scheduled Caste or Scheduled Tribe	0.17	0.02	4195
	[0.37]	(0.02)	
Other Backward Clasess	0.30	-0.04	4195
	[0.46]	(0.03)	
Economically Backward Classes	0.42	0.03	4195
	[0.49]	(0.03)	
Hindu	0.90	-0.01	4195
	[0.29]	(0.02)	
Muslim	0.10	0.01	4195
	[0.29]	(0.02)	
Married	0.87	-0.01	4195
	[0.33]	(0.01)	
Unmarried	0.03	0.01*	4195
	[0.16]	(0.01)	
No education	0.51	0.00	4195
	[0.50]	(0.02)	
Primary education (up to class 8)	0.30	-0.02	4195
	[0.46]	(0.01)	
Secondary or higher education	0.19	0.01	4195
	[0.40]	(0.01)	
Government connection	0.06	0.01	4195
	[0.24]	(0.01)	
Household members (includes the respondent)	6.15	0.15	4195
	[3.03]	(0.11)	
Lives in the location for more than 6 months a year	0.94	-0.00	4195
	[0.24]	(0.01)	
Number of assets owned (out of 12)	5.09	0.02	4195
	[1.85]	(0.07)	
Working	0.61	0.01	4195
	[0.49]	(0.02)	
Self-employed in agriculture	0.25	-0.01	4195
	[0.44]	(0.02)	

Notes: The first column shows the means of the outcome variable in the control group. The second column shows the regression coefficient of the outcome variable on an indicator variable for being treated, including block fixed effects. The sample is restricted to the beneficiary sample. *Government connection* measures social connections, it equals one if the respondent or a member of their family holds a panchayat/block position, and zero otherwise. Clustered standard errors at the panchayat level in parentheses. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table A6: First Stage – Blocks Where Audits Happened

	Full RCT Sample			Follow Up Survey Sample		
	Audited FY19-20	Number Audits FY21-22 Through FY23-24	Total Audits Through FY23-24	Audited FY19-20	Number Audits FY21-22 Through FY23-24	Total Audits Through FY23-24
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.659***	0.015	0.674***	0.723***	-0.000	0.723***
	(0.037)	(0.029)	(0.051)	(0.101)	(0.036)	(0.104)
Control Mean	0.03	2.08	2.12	0.01	2.18	2.19
N	2101	2101	2101	274	274	274

Notes: Bihar did not conduct standard social audits in FY20-21 due to the Covid-19 pandemic. All regressions are limited to panchayats in blocks where audits happened and include block fixed effects. Standard errors are in parentheses, clustered at the panchayat level. ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A7: Impact of Treatment on Administrative Expenditure on Citizen-Facing Benefits Under Workfare and Housing – Block-Specific Audit Phases

	Diff-in-Diff	ANCOVA	ANCOVA (M)	Diff-in-Diff	ANCOVA	ANCOVA (M)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Administrative Data - Workfare and Housing Labor Benefits						
	Household Demand Registered			Wage Expenditure		
Treatment	-0.42			-1.41		
	(3.20)			(7.61)		
Treatment × Anticipation Period	-0.43	-0.93	1.27	-1.56	-3.21	2.69
	(2.29)	(2.19)	(2.30)	(6.09)	(6.04)	(6.39)
Treatment × Post Period	-5.53**	-5.43*	-3.83	-16.11**	-15.82**	-11.73*
	(2.78)	(2.78)	(2.65)	(7.39)	(7.41)	(6.50)
Control Mean: Anticipation Period	116.20	116.20	116.20	293.05	293.05	293.05
Control Mean: Post Period	133.94	133.94	133.94	318.03	318.03	318.03
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512
Panel B: Administrative Data - Housing Registration and Materials Benefits						
	Housing Registrations			Housing Materials Expenditure		
Treatment	-23.25			-25.59**		
	(14.65)			(12.34)		
Treatment × Anticipation Period	-26.51	-19.39	-9.61	-2.22	11.18	10.83
	(40.26)	(40.45)	(38.77)	(21.83)	(20.88)	(20.97)
Treatment × Post Period	11.43	10.56	18.20	8.15	11.63	11.42
	(26.10)	(25.57)	(23.34)	(22.85)	(21.88)	(21.36)
Control Mean: Anticipation Period	38.16	38.16	38.16	331.34	331.34	331.34
Control Mean: Post Period	1096.51	1096.51	1096.51	957.19	957.19	957.19
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512

Notes: Column headers list outcome variables of regressions. The first and fourth columns run a differences-in-differences specification, including data from April 2018 until April 2021. The remaining columns use data from April 2019 until April 2021 and control for the pre-period (April 2018-March 2019) outcome mean. ANCOVA (M) specifications control for household labor demand registered, total wage expenditure, housing registriations, and housing materials expenditure. All expenditures are measured in ₹1,000s. The sample is restricted to blocks where at least one audit happened. All regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A8: Impact of Treatment on Administrative Expenditure on Citizen-Facing Benefits Under Workfare and Housing – Separating Implementation and Post Implementation Effects

	Diff-in-Diff	ANCOVA	ANCOVA (M)	Diff-in-Diff	ANCOVA	ANCOVA (M)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Administrative Data - Workfare and Housing Labor Benefits						
	Houshold Demand Registered			Wage Expenditure		
Treatment	-0.42			-1.41		
	(3.20)			(7.61)		
Treatment × Anticipation Period	3.74	3.67	5.41	14.30	14.03	18.56
	(4.10)	(4.16)	(4.19)	(13.86)	(14.26)	(14.51)
Treatment × Implementation Period	-3.42*	-3.49*	-1.75	-11.49**	-11.76***	-7.23*
	(2.00)	(1.86)	(1.98)	(4.54)	(4.04)	(4.33)
Treatment × Post-Imp. Period	-5.96*	-6.00*	-4.24	-16.86*	-17.06*	-12.46
	(3.28)	(3.32)	(3.12)	(9.38)	(9.45)	(8.37)
Control Mean: Anticipation Period	116.20	116.20	116.20	293.05	293.05	293.05
Control Mean: Implementation Period	94.88	94.88	94.88	215.06	215.06	215.06
Control Mean: Post-Imp. Period	164.04	164.04	164.04	397.35	397.35	397.35
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512
Panel B: Administrative Data - Housing Registration and Materials Benefits						
	Housing Registrations			Housing Materials Expenditure		
Treatment	-23.25			-25.59**		
	(14.65)			(12.34)		
Treatment × Anticipation Period	23.53	24.58	32.73*	-3.68	2.19	1.93
	(19.09)	(19.34)	(19.42)	(15.27)	(15.81)	(15.21)
Treatment × Implementation Period	15.02	16.06	24.22	3.51	9.38	9.12
	(34.88)	(34.93)	(32.80)	(20.36)	(20.04)	(19.53)
Treatment × Post-Imp. Period	-12.70	-11.22	-2.96	8.20	14.58	14.35
	(25.74)	(25.47)	(24.01)	(26.32)	(25.54)	(25.27)
Control Mean: Anticipation Period	38.16	38.16	38.16	331.34	331.34	331.34
Control Mean: Implementation Period	1531.31	1531.31	1531.31	905.42	905.42	905.42
Control Mean: Post-Imp. Period	761.61	761.61	761.61	997.06	997.06	997.06
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512

Notes: Column headers list outcome variables of regressions. The first and fourth columns run a differences-in-differences specification, including data from April 2018 until April 2021. The remaining columns use data from April 2019 until April 2021 and control for the pre-period (April 2018-March 2019) outcome mean. ANCOVA (M) specifications control for household labor demand registered, total wage expenditure, housing registriations, and housing materials expenditure. All expenditures are measured in ₹1,000s. The sample is restricted to blocks where at least one audit happened. All regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A9: Treatment on the Treated Effects: Impact of Treatment on Administrative Expenditure on Citizen-Facing Benefits Under Workfare and Housing

	Diff-in-Diff	ANCOVA	ANCOVA (M)	Diff-in-Diff	ANCOVA	ANCOVA (M)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Administrative Data - Workfare and Housing Labor Benefits						
	Household Demand Registered			Wage Expenditure		
Received an Audit	-0.64			-2.15		
	(4.86)			(11.53)		
Received an Audit × Anticipation Period	5.68	5.57	8.21	21.70	21.29	28.17
	(6.24)	(6.33)	(6.38)	(21.17)	(21.78)	(22.20)
Received an Audit × Post Period	-7.37**	-7.45**	-4.78	-22.04**	-22.39**	-15.45*
	(3.59)	(3.56)	(3.38)	(9.48)	(9.31)	(7.99)
Control Mean: Anticipation Period	116.20	116.20	116.20	293.05	293.05	293.05
Control Mean: Post Period	133.94	133.94	133.94	318.03	318.03	318.03
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512
Panel B: Administrative Data - Housing Registration and Materials Benefits						
	Housing Registrations			Housing Materials Expenditure		
Received an Audit	-35.28			-38.83**		
	(21.96)			(18.40)		
Received an Audit × Anticipation Period	35.71	37.29	49.66*	-5.58	3.31	2.92
	(28.56)	(28.94)	(28.91)	(23.21)	(23.97)	(23.05)
Received an Audit × Post Period	-0.97	0.98	13.44	9.36	18.68	18.32
	(39.16)	(38.98)	(35.80)	(31.50)	(30.41)	(29.77)
Control Mean: Anticipation Period	38.16	38.16	38.16	331.34	331.34	331.34
Control Mean: Post Period	1096.51	1096.51	1096.51	957.19	957.19	957.19
Number of Panchayats	2101	2101	2101	2101	2101	2101
N	77724	52512	52512	77724	52512	52512

Notes: Column headers list outcome variables of instrumental variables regressions where audit receipt (and its interactions with time period) are instrumented with treatment (and its interactions with time period). The first and fourth columns run a differences-in-differences specification, including data from April 2018 until April 2021. The remaining columns use data from April 2019 until April 2021 and control for the pre-period (April 2018-March 2019) outcome mean. ANCOVA (M) specifications control for household labor demand registered, total wage expenditure, housing registriations, and housing materials expenditure. All expenditures are measured in ₹1,000s. The sample is restricted to blocks where at least one audit happened. All regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A10: Impact of Treatment on Administrative Expenditure on Citizen-Facing Benefits Under Workfare and Housing – Blocks Where No Audits Happened

	Household Demand Registered			*Wage Expenditure*		
	Diff-in-Diff	ANCOVA	ANCOVA (M)	Diff-in-Diff	ANCOVA	ANCOVA (M)
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-6.904*			-13.030		
	(3.555)			(8.846)		
Treatment × Anticipation Period	1.188	-0.266	-0.337	-7.595	-10.660	-9.710
	(5.224)	(4.939)	(5.021)	(14.166)	(13.799)	(14.010)
Treatment × Post Period	5.865*	4.411	4.340	7.583	4.518	5.469
	(3.115)	(2.955)	(2.998)	(8.596)	(8.281)	(8.514)
Control Mean: Anticipation Period	117.09	117.09	117.09	278.62	278.62	278.62
Control Mean: Post Period	129.99	129.99	129.99	314.86	314.86	314.86
Number of Panchayats	1336	1336	1336	1336	1336	1336
N	49432	33400	33400	49432	33400	33400

Notes: The sample is restricted to blocks where no audits happened. The first and fourth columns run a differences-in-differences specification, including data from April 2018 until April 2021. The remaining columns use data from April 2019 until April 2021 and control for the pre-period (April 2018-March 2019) outcome mean. ANCOVA (M) specifications control for household labor demand registered, total wage expenditure, housing registriations, and housing materials expenditure. All expenditures are measured in ₹1,000s. All regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A11: Assessing Consequences of Declining Wage Expenditure Under Workfare – Treatment on the Treated Effects

	A. Decomposing Wage Expenditures			*B. Assessing Reallocation to Materials Expenditure*		
	Total Wage Expenditure	Workfare Wage Expenditure (₹1,000s)	Housing Wage Expenditure (₹1,000s)	Workfare Materials Expenditure (₹1,000s)	Total Wage Expenditure Share	Workfare Wage Expenditure Share
	(1)	(2)	(3)	(4)	(5)	(6)
Received an Audit × Anticipation Period	26.574	25.474	1.100	14.026	-0.010	-0.000
	(17.662)	(17.446)	(4.454)	(16.128)	(0.016)	(0.027)
Received an Audit × Post Period	-13.875*	-11.252*	-2.622	20.587	-0.020**	-0.045***
	(7.829)	(6.686)	(3.140)	(15.420)	(0.010)	(0.014)
Control Mean: Anticipation Period	232.45	157.02	75.43	97.87	0.84	0.68
Control Mean: Post Period	293.10	155.66	137.44	181.80	0.74	0.57
Number of Panchayats	2101	2101	2101	2101	2078	2064
N	52512	52512	52512	52512	6199	5694

Notes: Column headers list outcome variables of regressions. All regressions instrument audit receipt and its interactions with a treatment dummy and relevant interactions. The time period spans from April 2019 until April 2021. Material and wage expenditure is available at the panchayat × month level, but we calculate wage expenditure shares at the panchayat × time period level (pre, anticipation, implementation, and post-implementation period) to avoid missing values and smooth outliers driven by sporadic invoicing of materials expenditure. The sample is restricted to blocks where at least one audit happened. All regressions control for pre-period mean of the outcome variable (when available), alongside the mean pre-implementation housing material expenditure sanctioned, the mean pre-implementation housing material expenditure, the mean pre-implementation household demand registered, and the mean pre-implementation total wage expenditure (including workfare and housing). The total wage expenditure share divides total wage expenditure (for both workfare and housing) by the sum of total wage and workfare materials expenditure. The workfare wage expenditure share divides workfare wage expenditure by the sum of workfare wage ependiture and workfare materials expenditure. There were two panchayats that had a missing value for wage expenditure shares due to no expenditures in the pre period. We recode these values to the pre-period sample mean. Additionally, all regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A12: Treatment Effects on Number of Projects Supported by MGNREGS

	ANCOVA (M)		
	Total Projects (Workfare + Houses)	Workfare Projects	Houses
	(1)	(2)	(3)
Treatment × Anticipation Period	0.453	0.517	-0.064
	(1.079)	(0.387)	(0.992)
Treatment × Post Period	-0.294	0.040	-0.334
	(0.888)	(0.348)	(0.782)
Control Mean: Anticipation Period	31.02	3.76	27.26
Control Mean: Post Period	58.37	7.31	51.07
Number of Panchayats	2101	2101	2101
N	52512	52512	52512

Notes: Column headers list outcome variables of regressions. The time period spans from April 2019 until April 2021. The sample is restricted to blocks where at least one audit happened. All regressions control for the mean pre-implementation housing material expenditure sanctioned, the mean pre-implementation housing material expenditure, the mean pre-implementation household demand registered, and the mean pre-implementation total wage expenditure (including workfare and housing). Additionally, all regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A13: Impact of Audits on Program Graft and Misallocation – Treatment on the Treated Effects

		Index components				
	Misallocation index	Any Ghost Payment	Paid Kickback	Kickback	Payment Kept by Respondent (in Index Reversed)	Leakage
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Overall Effect						
Audit happened	-0.123**	-0.016	-0.020	-653.130**	2104.942	-3039.040**
	(0.052)	(0.023)	(0.020)	(321.400)	(1629.881)	(1537.943)
No Audit Mean	-0.03	0.43	0.24	2869.13	45329.50	13681.78
N	4341	4341	4193	4155	3999	3884
Panel B. By Beneficiary Type						
Audit Happened × Housing	-0.284***	-0.015	-0.079**	-1465.451**	4995.728	-6763.562*
	(0.101)	(0.026)	(0.039)	(718.750)	(3542.495)	(3490.714)
Audit Happened × Workfare	0.007	-0.018	0.027	5.145	-376.303	-68.272
	(0.051)	(0.037)	(0.018)	(169.800)	(763.776)	(913.801)
Housing	-0.148**	-0.575***	0.370***	5914.412***	87498.950***	22170.704***
	(0.058)	(0.025)	(0.021)	(401.800)	(1905.652)	(1940.888)
P-value: Audit × Workfare=Audit × Housing	0.01	0.95	0.01	0.05	0.15	0.07
No Audit Mean (Workfare)	0.07	0.71	0.06	105.09	770.49	4363.62
N	4341	4341	4193	4155	3999	3884

Notes: Column headers list outcome variables of instrumental variables regressions where audit receipt (and, in Panel B, its interactions with program type) are instrumented by treatment (and, in Panel B, its interactions with program type). All columns include outcomes from stacked data, where there is one observation per beneficiary for each program type as relevant. Panel A shows the overall effect. Panel B shows the decomposed effect by program type. The *misallocation index* is a GLS standardized index, which includes *any ghost payment*, *paid kickback*, *payment kept by respondent*, and *leakage* as components, with all components coded so higher values correspond to greater misallocation. Kickbacks, payment kept by respondent, and leakage are all measured in ₹s and topcoded at the 99th percentile. All regressions include block fixed effects and control for gender and a dummy to identify individuals in the housing sample. Details of variable construction are described in Appendix B. Clustered standard errors at the panchayat level. ***$p\leq0.01$, **$p\leq0.05$, *$p\leq0.1$.

Table A14: Impact of an audit on citizens' access to workfare – Treatment on the treated effects

	Participated in Workfare	Participated in or Paid for Workfare	Total Workfare Wages: Apr. 2019 - Mar. 2020
	(1)	(2)	(3)
Audit Happened	-0.035**	-0.036**	-81.552*
	(0.016)	(0.017)	(44.090)
No Audit Mean	0.055	0.066	114.105
N	1868	1868	1835

Notes: Column headers list outcome variables of instrumental variables regressions where audit receipt is instrumented with treatment. The sample is restricted to the voter roll sample. All regressions includes block fixed effects and control for gender. Total workfare wages is reported in ₹s and top-coded at the 99th percentile. Clustered standard errors at the panchayat level. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table A15: Assessing Consequences of Declining Wage Expenditure Under Workfare – Endline Survey Panchayats Only

	A. Decomposing Wage Expenditures			*B. Assessing Reallocation to Materials Expenditure*		
	Total Wage Expenditure	Workfare Wage Expenditure (₹1,000s)	Housing Wage Expenditure (₹1,000s)	Workfare Materials Expenditure (₹1,000s)	Total Wage Expenditure Share	Workfare Wage Expenditure Share
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment × Anticipation Period	-4.147	-0.912	-3.236	6.328	0.022	0.061
	(37.091)	(38.944)	(6.653)	(29.213)	(0.027)	(0.038)
Treatment × Post Period	-40.788***	-36.359***	-4.428	-5.663	-0.005	-0.032
	(11.778)	(10.929)	(5.248)	(24.683)	(0.019)	(0.027)
Control Mean: Anticipation Period	233.43	172.42	61.01	79.18	0.87	0.74
Control Mean: Post Period	274.61	140.95	133.66	176.60	0.73	0.55
Number of Panchayats	274	274	274	274	274	270
N	6850	6850	6850	6850	816	717

Notes: Column headers list outcome variables of regressions. The time period spans from April 2019 until April 2021. Material and wage expenditure is available at the panchayat × month level, but we calculate wage expenditure shares at the panchayat × time period level (pre, anticipation, implementation, and post-implementation period) to avoid missing values and smooth outliers driven by sporadic invoicing of materials expenditure. The sample is restricted to blocks where at least one audit happened. All regressions control for pre-period mean of the outcome variable (when available), alongside the mean pre-implementation housing material expenditure sanctioned, the mean pre-implementation housing material expenditure, the mean pre-implementation household demand registered, and the mean pre-implementation total wage expenditure (including workfare and housing). The total wage expenditure share divides total wage expenditure (for both workfare and housing) by the sum of total wage and workfare materials expenditure. The workfare wage expenditure share divides workfare wage expenditure by the sum of workfare wage ependiture and workfare materials expenditure. There were two panchayats that had a missing value for wage expenditure shares due to no expenditures in the pre period. We recode these values to the pre-period sample mean. Additionally, all regressions include fixed effects at the block-month level. Clustered standard errors at the block level in parentheses: ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table A16: Impact of treatment on components of the workfare awareness index

	Aware of Workfare	Aware of Job Card	Aware of Days Entitled to Work	Aware of Payment Method	Aware of Daily Wage	Aware of Eligibility to Work
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Overall Effect						
Treatment	0.010	0.004	-0.002	0.014	-0.010	-0.000
	(0.008)	(0.014)	(0.005)	(0.016)	(0.008)	(0.017)
Control Mean	0.91	0.77	0.04	0.49	0.06	0.64
N	6278	6278	5810	5833	5810	5819
Panel B. By Beneficiary Type						
Treatment × Housing	0.008	0.012	0.008	0.031	-0.006	0.016
	(0.016)	(0.022)	(0.007)	(0.024)	(0.010)	(0.026)
Treatment × Workfare	0.011	0.006	-0.009	0.003	-0.019	-0.031
	(0.010)	(0.016)	(0.010)	(0.025)	(0.016)	(0.022)
Treatment × Voter Roll	0.011	-0.005	-0.005	0.011	-0.003	0.023
	(0.015)	(0.024)	(0.008)	(0.025)	(0.009)	(0.028)
Housing	-0.039***	-0.071***	-0.036***	-0.171***	-0.057***	-0.093***
	(0.011)	(0.017)	(0.009)	(0.025)	(0.015)	(0.021)
Voter Roll	-0.049***	-0.120***	-0.033***	-0.133***	-0.062***	-0.122***
	(0.013)	(0.019)	(0.009)	(0.024)	(0.013)	(0.022)
P-value: Treat × Workfare= Treat × Housing	0.88	0.79	0.18	0.39	0.50	0.11
Control Mean (Workfare)	0.94	0.84	0.07	0.59	0.10	0.71
N	6278	6278	5810	5833	5810	5819

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. All columns include outcomes from stacked data, where there is one observation per program for each beneficiary type as relevant. Panel A shows the overall effect. Meanwhile, Panel B shows the decomposed effect by beneficiary type. All regressions include block fixed effects and control for gender and individuals drawn from the housing sample. Clustered standard errors at the panchayat level. ***$p\leq0.01$, **$p\leq0.05$, *$p\leq0.1$.

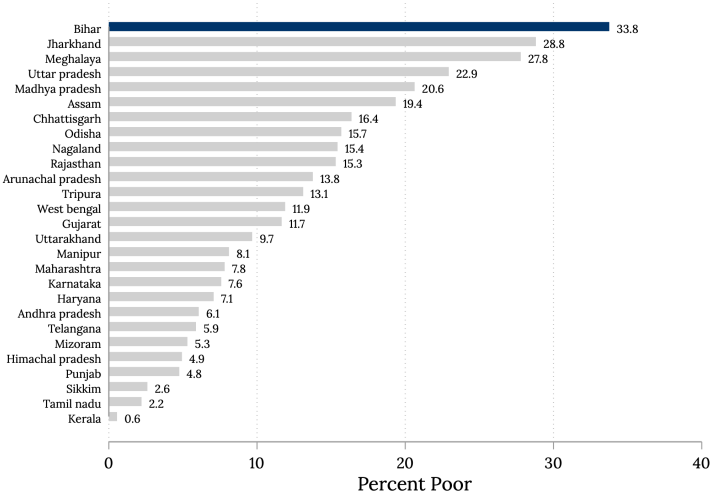
Table A17: Impact of treatment on audit community meeting participation

	Remembers Community Meeting (1)	Attended Community Meeting (2)
Panel A. Overall Effect		
Treatment	0.050***	0.022***
	(0.014)	(0.008)
Control Mean (Workfare)	0.29	0.11
N	6254	6243
Panel B. By Beneficiary Type		
Treatment × Housing	0.065***	0.030**
	(0.020)	(0.013)
Treatment × Workfare	0.053**	0.028*
	(0.023)	(0.015)
Treatment × Voter Roll	0.032	0.004
	(0.020)	(0.011)
Housing	-0.044**	-0.037***
	(0.019)	(0.013)
Voter Roll	-0.080***	-0.048***
	(0.020)	(0.013)
Control Mean (Workfare)	0.29	0.11
N	6254	6243

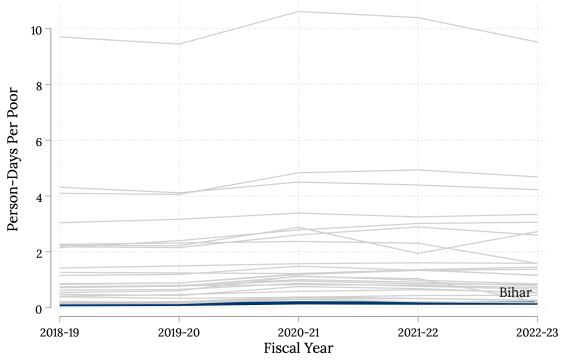
Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. All columns include outcomes from stacked data, where there is only one observation for program type. Panel A shows the overall effect. Meanwhile, Panel B shows the decomposed effect by program type. All regressions include block fixed effects and control for gender, individuals drawn from the housing sample and invididuals drawn from the voter roll sample. Clustered standard errors at the panchayat level. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Figure A1: Poverty and Implementation of Workfare and Housing Among Indian States

A. Poverty Rate by State

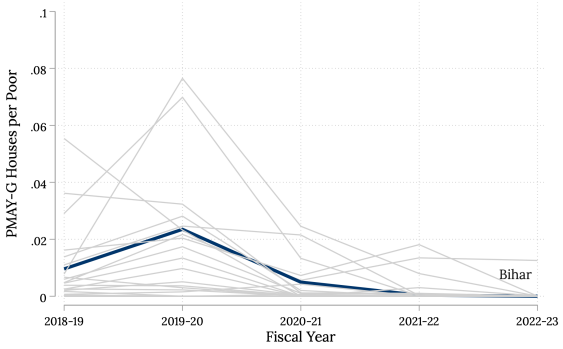


B. MGNREGS Person-Days Per Rural Poor



Notes: Blue line indicates the state of Bihar and gray lines show the other states of India.

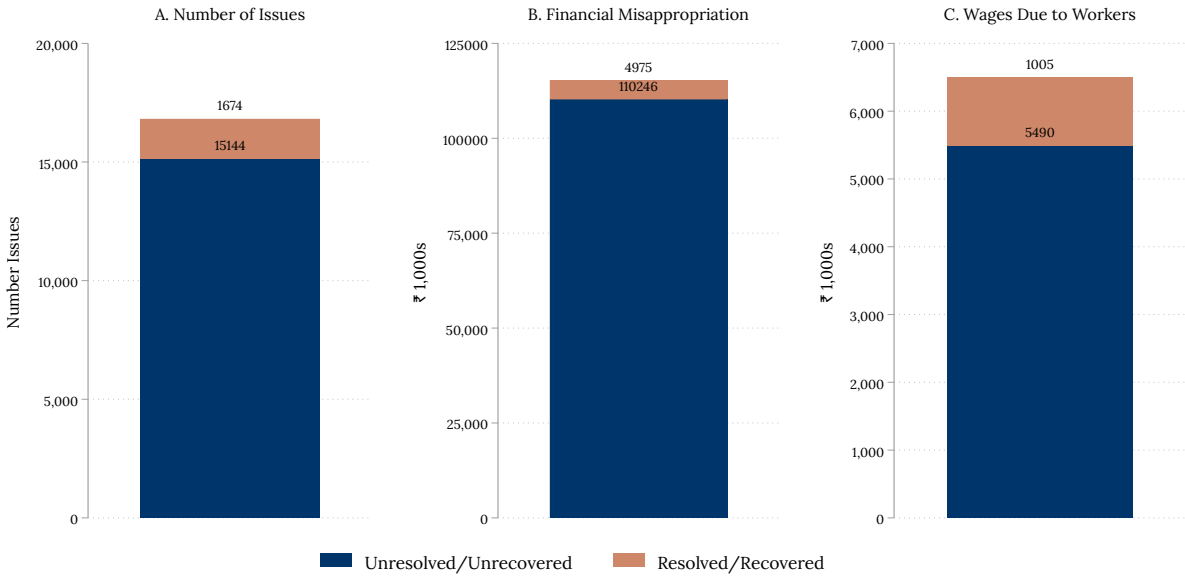
C. PMAY-G Homes Completed Per Rural Poor



Notes: Blue line indicates the state of Bihar and gray lines show the other states of India. Graph drops Tripura for readability.

Notes: Panel A shows the poverty rate in 27 states in India. Data come from Reserve Bank of India's Handbook of Statistics on Indian States. We report the percent poor per the multidimensional poverty index constructed from the 2019-21 National Family Health Survey. Panel B indicates MGNREGS person-days per poor across 27 states in India. We calculate the number poor by using the poverty rate in Panel A and 2023 population estimates from the Ministry of Health and Family Welfare's Technical Group on Population Projections. Panel C shows the number of completed PMAY-G homes per poor across 26 states in India (for readability we drop Tripura). MGNREGS data come from MGNREGS MIS Table R5.1.1 Employment Generated and PMAY-G homes data come from MIS Table 6.9. Sanctioned PMAY-G Houses and Mandays Generated. We assume zero homes were constructed under PMAY-G when states are not listed on the relevant MIS table. In India, a fiscal year starts on the 1st of April every year, and ends on the 31st of March.

Figure A2: Issue Identification and Resolution in Social Audits; FY 2019-2020



Notes: Data retrieved from the MGNREGS MIS, Table R9.3.1 – Social Audit Action Taken Report. Table accessed Sept. 7, 2023.

Figure A3: Timeline

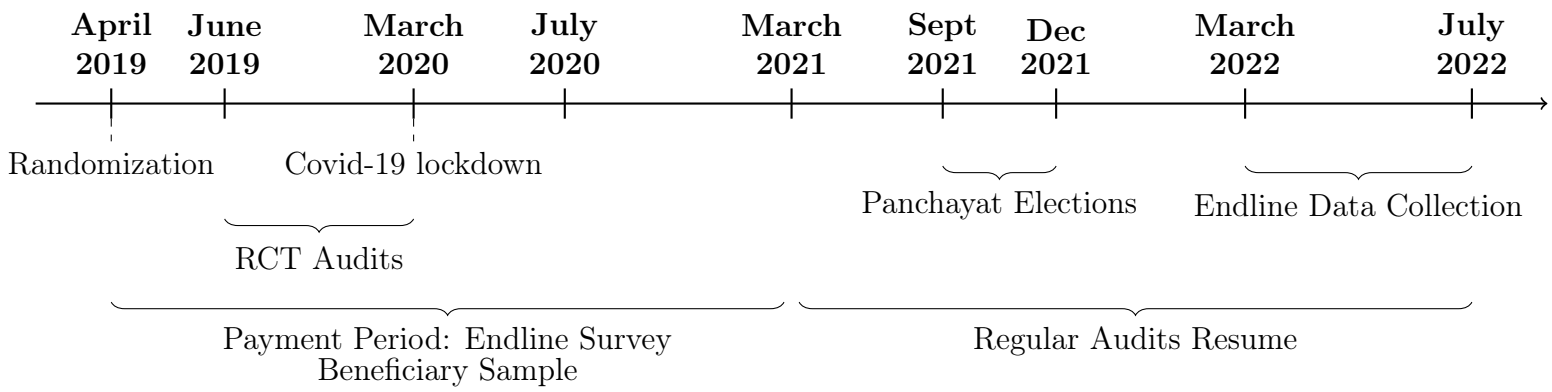


Figure A4: Map of RCT Blocks

Study & non-Study blocks in Bihar

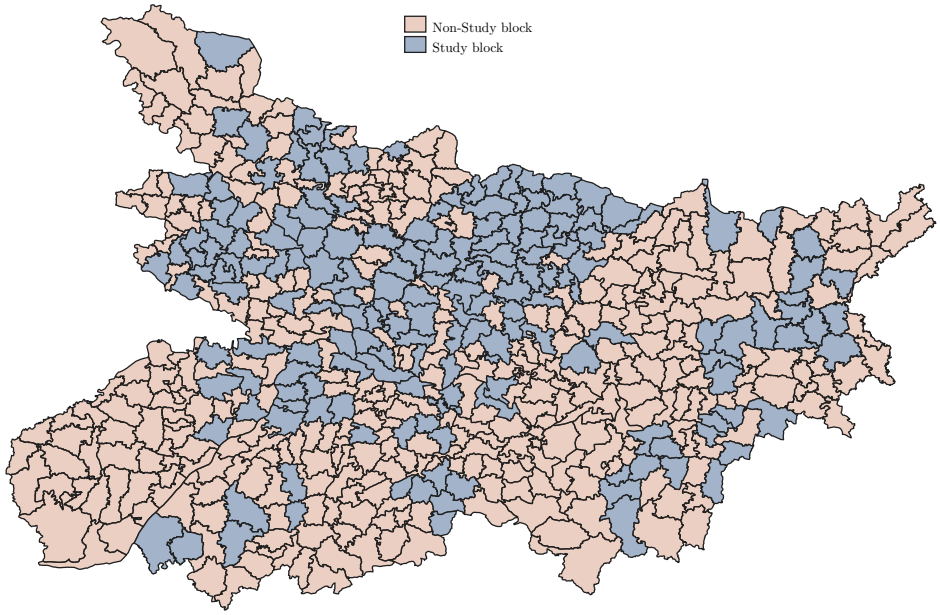
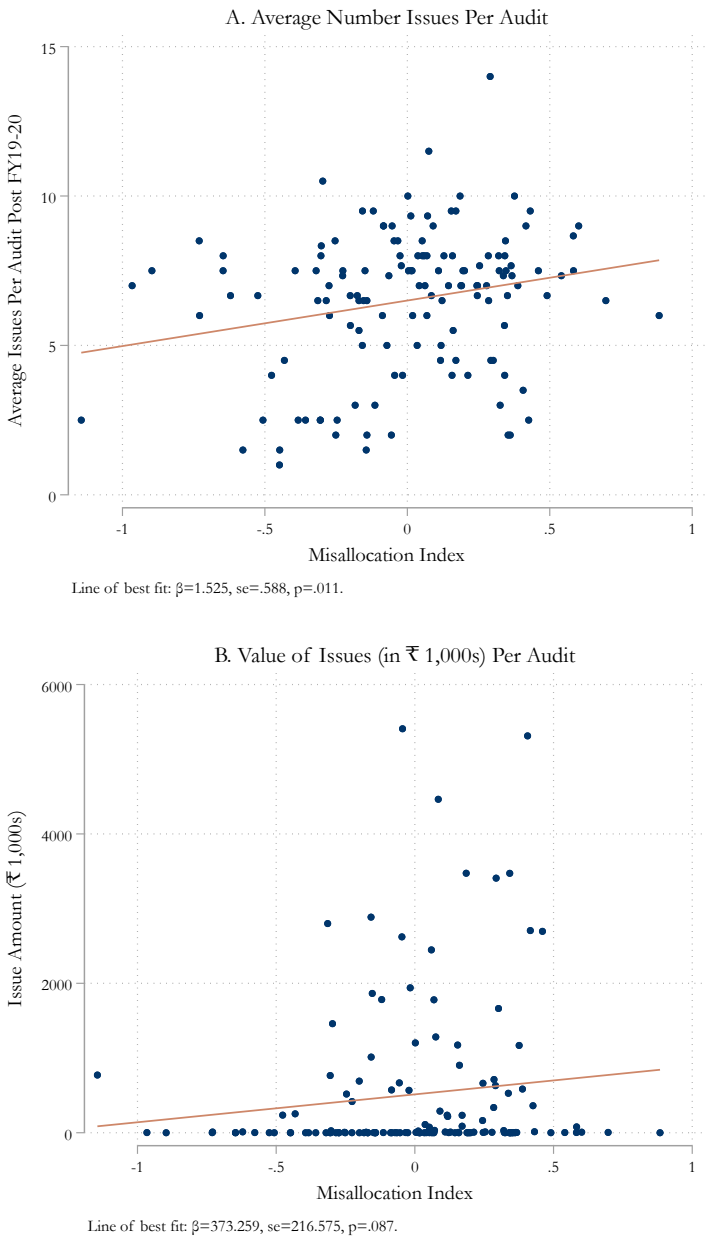


Figure A5: Correlation Between Survey Measures of Graft (Misallocation Index) and Post FY 2019-20 Audit Findings



Notes: This figure plots the correlation between panchayat-level averages of the misallocation index (x-axis), and the average number of issues per audit (Panel A) and total value of issues per audit (Panel B) in post FY 2019-20 audits. Notes to each figure report coefficients and standard errors (and associated p-values for a test of $\beta = 0$) for the line of best fit, using heteroskedastic standard errors. We limit the sample to control group panchayats with endline survey and post-FY 2019-20 audit data. N=136 in both panels.

B Data and Variable Construction

This appendix provides detail on data sources and how key variables in our analysis were calculated.

B.1 MGNREGS Administrative Data

In order to understand the impact of audits on overall MGNREGS expenditure, we study the following variables:

- *Household demand registered* – the total number of households for whom workfare demand was formally registered (per panchayat and month). Scraped from the MGNREGS Public Data Portal and available from April 2018-April 2021..
- *Wage expenditure* – total MGNREGS wage expenditures per panchayat, per month. This variable is constructed using information on the number of person-days (per panchayat per month) associated with workfare and housing. To construct a meansure of wage expenditure, we multiply the number of person-days by the MGNREGS for the relevant fiscal year. Scraped from the MGNREGS Public Data Portal and available from April 2018-April 2021.
- *MGNREGS materials expenditure* – the total amount of material expenditure billed per panchayat per month. We scrape this from the MGNREGS Materials Register in the MIS and use information on bill dates, aggregating up to the panchayat-month level. In months where there is no record of billed expenditures, we assign zero values. Data are available from April 2018 to March 2021.

To assess the separate impacts of audits on workfare and housing wage expenditures, we rely on two databases available at the work roll level, which we accessed via an API (not publicly available). The first database, covering April 2018 to April 2021, reports information on the number of person-days of employment provided, work end date, and work start date. The second, spanning April 2019 to April 2021, contains a code indicating the type of work, clearly distinguishing between housing linked payments and other projects. We merge these two datasets together to differentiate between work type, and multiple person-days by the MGNREGS wage for the relevant fiscal year to calculate wage expenditure.

- *Housing wage expenditure* – the total amount of MGNREGS wage expenditure paid to support construction of PMAY-G homes. Available at the panchayat-month level from April 2019 to April 2021.
- *Workfare wage expenditure* – the total mount of MGNREGS wage expenditure netting out housing wage expenditure. Available at the panchayat-month level from April 2019 to April 2021.

- *MGNREGS wage expenditure* – the sum of workfare and housing-linked expenditures using the above-referenced datasets; an alternative to the "wage expenditure" variable calculated using data scraped from the public data portal. Available at the panchayat-month level from April 2019 to April 2021.

- *MGREGS wage expenditure share* – MGNREGS wage expenditures (including both housing and workfare-linked payments) divided by the sum of total MGNREGS wage and MGNREGS materials expenditures. Since there are zero wage and materials expenditures in many months, and since materials expenditures are invoiced more sporadically than wage expenditures, we construct this variable at the panchayat by time period (pre-RCT, anticipation, implementation, and post-implementation period) level instead of by panchayat-month.

- *Workfare wage expenditure share* – workfare wage expenditures divided by the sum of workfare wage and MGNREGS materials expenditures. Since there are zero wage and materials expenditures in many months, and since materials expenditures are invoiced more sporadically than wage expenditures, we construct this variable at the panchayat by time period (pre-RCT, anticipation, implementation, and post-implementation period) level instead of by panchayat-month.

B.2 PMAY-G Administrative Data

- *Housing registrations* – the total number of new PMAY-G homes approved to begin construction (per panchayat per month). Scraped from the PMAY-G MIS and available from April 2018 to April 2021. These data were only scraped for RCT blocks where at least one audit happened.

- *Housing materials expenditure* – the total amount of expenditure on housing payments made through the PMAY-G system (per panchayat per month). These payments are meant to cover the materials needed to construct the beneficiary's home, but use of these funds is left to the discretion of the beneficiary. Scraped from the PMAY-G MIS and available from April 2018 to April 2021. These data were only scraped for RCT blocks where at least one audit happened.

B.3 Survey Data

B.3.1 Sampling Panchayats

We aimed to balance power and representativeness when selecting panchayats for the survey sample. We first identified a set of districts that were (a) logistically feasible for in-person surveys, (b) exhibited administrative data treatment effects in line with our statewide results, and (c) had sufficient blocks where at least 35 percent of panchayats selected for audit were actually audited.

After districts had been selected, we limited attention to blocks that had at least 7 treatment and 7 control panchayats, enrolling an equal number in each block until we reached our quota of 274 panchayats. We endeavored to enroll the maximum number of panchayats per block possible. Thus if a block had, for example, 8 control and 10 treatment panchayats, we would enroll all 8 controls and randomly select 8 of the 10 treatment panchayats for enrollment.

B.3.2 Sampling Respondents

Our survey data include two sub-samples: first, a sample of individuals who, per administrative data, received either workfare or housing-linked MGNREGS payments between April 2019 and April 2021. These individuals were asked details about whether they received workfare (for those listed as receiving workfare-linked payments) or PMAY-G (for those listed as receiving housing-linked payments). For workfare, we leveraged detailed worksite project descriptions in the administrative data to identify distinct MGNREGS projects (e.g. building a village cowshed, repairing a specified road) during our observation period. We used payment data to extract the dates an individual was recorded as working on a project and use project information plus payment data to define a "work engagement" – a period of work on a given project allowing for temporary breaks of up to 14 days on the project. We asked respondents to verify up to 4 work engagements, prompting them with information on the project and duration of the work. Roughly 90 percent of workers had 4 work engagements or less. For those who had more than 4, we sampled 4 at random. For housing, we asked individuals to report on total funds received for a PMAY-G house. Individuals who were recorded as benefiting from both workfare and housing were asked about participation in and receipt of benefits in both programs. For such individuals (N=127), we include separate observations for each program in our dataset. Individuals in the voter roll sample were asked to report workfare participation on a quarterly basis over our reference period.

B.3.3 Variables

Program Graft and Misallocation

Outcomes related to program graft and misallocation are available for individuals sampled from administrative MGNREGS data, including both workfare and housing beneficiaries.

- *Any ghost payment*
  - *Workfare beneficiaries* - This variable equals to one if the respondent reported that they did *not* work on one (or more) work engagements identified in the administrative data.
  - *Housing beneficiaries* - This variable equals one if the respondent reports that the did not build a PMAY-G home during the period for which we observe housing-linked MGNREGS payments.

- *Paid kickback*
  - *Workfare beneficiaries* - based on the question "Did you have to pay part of the money you got into your account to someone else?". This variable equals one if the respondent had to pay someone else workfare wages paid into their account, and zero otherwise.
  - *Housing beneficiaries* - based on the question "Did you have to pay part of this money to someone else? Please don't include any payments made for the construction of your house, such as payments to laborers, or to shopkeepers for materials". This variable equals to one if the respondent had to pay someone else, and zero otherwise.
- *Kickback*
  - *Workfare beneficiaries* - this is the total amount of workfare wages the respondent had to pay to someone else from their bank account associated with sampled work engagements. This variable istop-coded at the 99th percentile.
  - *Housing beneficiaries* - this is the total amount of money received under PMAY-G that the respondent had to pay to someone else, excluding valid expenditures for labor and materials. This variable is top-coded at the 99th percentile.
- *Payment kept by the respondent*
  - *Workfare beneficiaries* - equal to the total amount of money the respondent reported receiving in their bank account minus any kickbacks. For those who reported getting no payment, this variable equals zero. This variable is top-coded at the 99th percentile.
  - *Housing beneficiaries* - equal to the total amount of money the respondent reported receiving for their PMAY-G home (encompassing materials and labor compensation) minus any kickback. For those who reported no payment, this variable is coded to zero. This variable is top-coded at the 99th percentile.
- *Leakage* - this is the difference between the amount sent to the respondent per administrative data and the amount kept by the respondent (netting out for kickbacks). This variable is top-coded at the 1st and 99th percentile.
- *Misallocation index* - this is a GLS-weighted standardized index (following (Anderson, 2008)), aggregating *any ghost payment*, *paid kickback*, *kickback*, *payment kept by the respondent* and *leakage*. We flip the sign of *payment kept by the respondent* so that less beneficial outcomes always contribute to a higher score.
- *Paid kickback to elected official* - this is a dummy variable equal to one if the respondent reported paying a kickback *and* reported that some of the kickback went to an elected official. Elected officials include the panchayat head (mukhiya or sarpanch), ward member, or panchayat committee member

- *Paid kickback to contracted official* - this is a dummy variable equal to one if the respondent reported paying a kickback *and* reported that some of the kickback went to a contracted official. Contracted officials include the panchayat sachiv, the panchayar rozgar sahayak, the awaas sahayak, an MGNREGS mate, and block-level officials.

Access to Workfare

To study impacts on access to workfare in the general population, we use the voter roll sample.

- *Participated in workfare*
  - *Voter roll sample* - based on "Did you perform any work under MGNREGS in the period from April 2019 to March 2021?". It equals to one if the respondent confirmed working for MGRNEGS, and zero otherwise.
- *Participated in or paid for workfare*
  - *Voter roll sample* - this dummy variable is equal to one for anyone who reported participating in workfare *and* anyone who reported receiving payments without working.
- *Total workfare wages: April 2019 - March 2020*
  - *Voter roll sample* - the total amount amount of money received under workfare between April 2019 and March 2021. This variable is top-coded at the 99th percentile.

Demand for and Awareness of Workfare

- *Workfare reservation wage* - this is derived from an unfolding bracket question, which asked about the respondent's willingness to work under workfare. Respondents are first given a detailed description of workfare, using pictures to ensure they understand the nature of the work. We assign a reservation wage equals to ₹550 to any individual who stated they would either require a higher wage than ₹550 or would not work for workfare for any wage.
- *Workfare demand* - derived from "Do you want to do manual labour for a wage of 210 rupees, under MGNREGS? We would like to remind you that this work will be within your Panchayat or somewhere nearby. We would also like to remind you that the wages for this work will get credited to your bank account at least 15 days after completion of work. If you tell us that you would like to work, then we will share the information received from you with the Panchayat and Block-level officials, and they might contact you accordingly." This variable is a count of the total number of household members (including the respondent) for which the respondent chooses to submit workfare demand.

- *Workfare awareness index* - a GLS-weighted average (following (Anderson, 2008)) of six components — *aware of workfare*, *aware of job card*, *aware of days entitled to work*. *aware of payment method*, *aware of daily wage*, *aware of eligibility to work* (following (Kling et al., 2007)). Each index component is calculated by subtracting the control group mean and dividing by the control group standard deviation. The final index sums up all standardized index components and then takes the average.

- *Aware of workfare* - equal to one if the respondent answered "yes" to the following question "Have you heard of MGNREGS, which is a government scheme where each rural household can get manual labor work against daily wage payment. You may also know this as "rozgar guarantee" or "mitti ka kaam". The work available under MGNREGS is manual labour, used in tasks such as digging canals, laying roads, building ponds, planting trees, etc.". For those who replied "no" or "don't know", this variable equals to zero.

- *Aware of job card* - derived from "Do you know about a document called a job card? It mentions details of all eligible members of the household. If any member of the household work, that information is registered in the job card. The main page of the job card looks like this [show picture attached]". For those who replied "no" or "don't know", this variable equals zero.

- *Aware of days entitled to work* - equal to one if the respondent answered "100 days" to the question "How many days of work is the household entitled to get in a year as per the rules of MGNREGS". For those who answered a different number of days or didn't know the answer, this variable equals zero.

- *Aware of payment method* - equal to one if the respondent answered "Paid directly to their own account" to the multiple choice question "When people work for MGNREGS, how should they be paid - what are the rules about this". For those who replied something else or didn't know the answer, this variable equals zero.

- *Aware of daily wage* - equal to one if the respondent answered "198 rupees" ($\pm$10 rupees) or "210 rupees" ($\pm$10 rupees) to the question "According to the rules, what is the daily wage people should get when they work for MGNREGS?", zero otherwise. We count two values as correct because the daily wage changed mid survey (from 198 to 210).

- *Aware of eligibility to work* - based on "Do you think you are eligible, as per MGNREGS rules, to find work under MGNREGS?". This variable equals to zero if the respondent didn't know the answer.

Community Meeting (Gram Sabha) Participation

- *Remembers community meeting* - derived from "Between April 2019 and March 2021, has there ever been a ward sabha or gram sabha in your village?". If the respondent didn't know

the answer, this variable equals zero.

- *Attended gram sabha* - based on "Did you ever attend any of these meetings [gram sabha]?". For those who didn't remember a gram sabha, this variable equals zero.

Satisfaction with Local Leaders

- *Thinks local leader can help in receiving benefits* - respondents were asked "In general, do you think your local leaders are able to help you receive benefits from government programs?". The possible responses include "yes", "yes, but they help in exchange of some money/benefits" and "no". This variable is coded to 1 if they select "yes", and zero otherwise.

B.4 Socio Economic and Caste Census (SECC)

To explore pre-audit differences across panchayats, we utilize the 2011 SECC data that provides information on the socio-economic status and caste demographics of Indian households. We are not able to match all panchayats in our RCT sample to the SECC due to changes in names and panchayat boundaries.

- *Percent households economically marginalized* - this is the percent of households in the panchayat that have at least one SECC deprivation or automatically classified as poor. The SECC deprivations are as follows:
  - Households with only one room, kucha (temporary) walls and kucha (temporary) roof
  - No adult member between the ages of 16 and 59
  - Female headed households with no adult male member between 16 and 59
  - Households with disabled member and no able bodied adult member
  - scheduled caste or tribe (SC/ST) households
  - Households with no literate adult above 25 years
  - Landless households deriving a major part of their income from manual casual labor

  The criteria for automatic classification as poor are:
  - Unsheltered
  - Destitute
  - Living on alms
  - Manual scavenger families
  - Primitive tribal groups
  - Legally released bonded labor

We do not detail other variables here (e.g. total population, percent female) as they are self-explanatory.

B.5 Electoral Data

In order to assess audits' impacts on electoral outcomes, we collected results for mukhiya (panchayat leader) elections in 2016 and 2021. The 2016 data include name, gender, and party affiliation of winners. The 2021 data include name, gender, and party affiliation of all who stood for election. We merge based on panchayat and candidate name across the two to identify incumbents and challengers in 2021. We were able to complete this merge for 1,788 RCT panchayats in blocks where audits happened (others could not be merged due to changes in panchayat names and boundaries).

- *Incumbent ran* - a dummy equal to 1 if the 2016 winner stood for election in 2021 and zero otherwise.
- *Number challengers* - the number of non-incumbent challengers who stood for election in 2021.
- *Incumbent won* - a dummy identifying panchayats where the incumbent won. Equal to zero if the incumbent did not stand for election
- *Incumbent vote share* - the share of total votes in the panchayat captured by the incumbent in 2021.

B.6 Process Monitoring Data

To understand how audits were actually conducted and the extent to which they met their stated goals, we conducted a *process monitoring exercise*, which included extensive data collection across 26 GPs in 13 districts and was designed to identify ways the state could improve auditing while clarifying unique challenges and benefits associated with audits utilizing female auditors.

This exercise consisted of three main components: shadowing auditors (called social audit resource persons or SARPs) to evaluate their adherence to audits standards and protocols; interviewing SARPs and other stakeholders in the process to evaluate challenges and better understand experiences with audits; and conducting our own labor verification exercise with random subset of "listed workfare workers" (individuals who appeared in official administrative data as having worked for workfare in FY 2018-19) to evaluate how comprehensive current audits are.

Our team shadowed SARPs while they conducted the labor verification process in 380 households and shadowed SARPs when conducting asset verification of 24 worksites. In addition, we spoke with 46 SARPs, 42 other audit staff members (including Team Leaders and District Resource Persons, or DRPs), 26 elected panchayat leaders (mukhiyas), and 26 PRSes, the local MGRNEGS implementing bureaucrat. We independently surveyed 362 listed workers to understand their experiences with

workfare and the audit. In half the panchayats we surveyed workers after the community meeting (gram sabha), to measure attendance, while in the other half we surveyed workers before (to measure intent to attend). Enumerators also observed 13 community meetings, coding attendance and other process outcomes. We do not provide additional detail on variable construction here, as it is self-explanatory based on naming conventions.

B.7 Audit Reports Data

In order to understand both the issues identified in study audits (conducted in FY 2019-2020) as well as issues identified in subsequent audits, we scraped social audit reports from the MGNREGS MIS portal. Data are available for FY 2019-20, 2021-22, 2022-23, and 2023-24. There are no audit reports for FY 2020-21 because the usual audit process was put on hold due to Covid-related restrictions. These data include a list of issues identified during the audit, and issue classification (i.e. financial misappropriation, financial deviation, process violation, grievance). Issues amounts are also available when relevant (not all issues map to an issue amount) We use these data to construct the following analysis variables:

- *Total issues post FY19-20* - the total number of issues (per panchayat) identified in audits conducted from FY 2021-22 to FY 2023-24.
- *Total issues pre panchayat elections* - includes issues identified in audits conducted in FY 2021-22.
- *Total issues post panchayat elections* - includes issues identified in audits conducted in FYs 2022-23 and 2023-24.
- *Value of issues post FY19-20 variables* - analogous to the total issues variables above, but instead records the total value of issues in ₹.

We also use data from the social audits "action taken report" (available on the MGNREGS MIS) to characterize the number of issues that were (a) identified and resolved versus unresolved; (b) the amount of financial misappropriation unrecovered and recovered; (c) wages due to workers recovered and unrecovered.