The Value of Decentralized Decision-Making: Lessons from an Agricultural Infrastructure Experiment

PRELIMINARY. DO NOT CITE OR CIRCULATE

Xavier Giné, Aprajit Mahajan, Anup Malani, Manaswini Rao*

January 30, 2025

Abstract

We study barriers to collective action in the provision of common-pool resources by examining surface irrigation in rural India. Across 240 villages in Telangana, we randomize whether a village receives a field channel construction intervention or a budget-equivalent untied transfer. 20% of the grant villages constructed a field channel, while 99% of the channel intervention villages did so. The constructed channels increase irrigation (by 40 days, 45% or 0.4 s.d.), yields (0.17 s.d.) and net revenues (0.08 s.d.). We examine the role of program design in collective action problems by experimentally varying: (a) choice over channel location which contrasts a one-person-one-vote referendum to a data-driven benchmark

*Giné: The World Bank DECRG (xgine@worldbank.org), Mahajan: UC Berkeley, Department of Agricultural and Resource Economics (aprajit@berkeley.edu), Malani: UChicago, The Law School (amalani@uchicago.edu), Rao: U Delaware, Department of Economics (manaswini.rao@gmail.com). We thank Francisco Costa, Olga Gorbachev, Adrienne Lucas, Parag Mahajan, Jeremey Tobacman, and seminar participants at the UDel Economics seminar. Nitesh Bharadwaj, Nidhila Adusumalli, Rohini M S, Sandhya Seetharaman, Vrinda Kapoor, and Neeraja Ali provided excellent research support. This study was funded through generous grants from J-PAL (GI and K-CAI funding initiatives), FID, and the World Bank. This study was preregistered with the AEA RCT Registry under ID Number AEARCTR-0004336. Research was conducted with approval from the Institutional Review Boards at the University of Chicago (IRB16-1404-AM008) and at the Institute for Financial Management and Research in Chennai, India (IRB00007107). Views expressed in this paper are those of the authors, and do not necessarily reflect either the opinions of the World Bank, its executive directors, the countries they represent, or the views of the federal government of India. The authors have no financial interests to disclose.

based on private preference elicitation, and (b) choice over channel construction that compares execution by local community members to that by an external third-party contractor. The location referendum results diverged from the data-driven benchmark and lowered irrigation by 10 days (25%). The referendum also distributed irrigation resources towards more elite farmers. Lastly, there are no significant differences between community and third-party constructed projects. Taken together, these findings suggest that in our context the data-driven approach with private information outperforms more participatory alternatives.

1 Introduction

Investing in infrastructure requires collective action and coordination between multiple stakeholders. On the one hand, standard textbook solutions to overcoming coordination problems include common ownership or centralized provision. However, local communities may have divergent preferences and ultimately bear the cost. Given this tension, what is the best way to give local communities agency in designing infrastructure projects? What are the costs of participation? These are open questions and the literature on decentralization is unclear on when and why decentralization generates socially optimal outcomes (Hayek, 1945; Mookherjee, 2015).[1]

In this paper, we examine the presence or absence of collective action and coordination failures in the context of surface irrigation infrastructure. We further present experimental evidence of different degrees of community involvement on the aggregate benefits of infrastructure project choice, quality of project implementation, ongoing maintenance, and economic productivity in a context where preferences over the project - location of the irrigation infrastructure - are well defined and can be elicited ex-ante. The specific infrastructure project that we examine is the construction of field surface irrigation channels linking local irrigation reservoirs (tanks) to agricultural plots in the downstream command area in 240 villages in Telangana, India, using a randomized controlled trial design. Channel construction requires collective action

[1] For example, centralized provision of infrastructure such as roads, irrigation canals, etc., can affect local communities in both positive and negative ways. Centralized provision could resolve the coordination problem but information problems could driven a wedge between community and central agent's incentives. In contrast, a decentralized approach requiring community members to participate in programs to identify locally suited strategies may resolve the information problem but could reintroduce coordination problem while also incurring substantial time costs for the community members in participating in such programs, often involving voluntary time and labor contributions. In communities with a significant number of poor, such costs impose a regressive tax.

among farmers in the command area, where farmers face threats of free-riding behavior and require cooperation from other farmers who need to allow water to pass downstream.

At baseline, we collected rich data on irrigation availability, agricultural productivity, production costs, and private willingness to pay for each feasible location of irrigation channels among a representative sample of more than 6200 constituent farmers. We find that around 45% of the irrigation channels in the study villages were in disrepair and not used, but farmers on average would have derived large benefits through additional revenue if these channels were functional. This raises the question whether an external intervention could solve this coordination problem. Alternatively, the status of disrepair could reflect financial constraints affecting consumption smoothing, where even though farmers derived benefits at the end of the agricultural season, upfront costs of construction prevented the provision and upkeep of surface irrigation infrastructure.

This paper speaks to the core discussion in the field of decentralization of public good provision and community-driven-development. The argument for decentralization is based on instrumental and intrinsic values of decentralization (Balán et al. 2022; Bardhan and Mookherjee 2006; Dal Bó et al. 2010; Ostrom 1990), with instrumental value from better information on local preferences and improved fund monitoring (Balán et al. 2022; Chambers 1984; Oates 1999, 1972), and intrinsic value from empowering communities and strengthening their capacity for program implementation (Dal Bó et al. 2010; Ostrom and Gardner 1993). However, existing literature on the effects of decentralization and community participation suffers from a key challenge: comparison of final welfare outcomes depends on the endogenous choice over the type of good (for example, a school vs. village road), making it hard to rule out the role of program design from the choice of the good itself. Despite the lack of robust evidence for or against the role of the community in development programs (Casey 2018; Mansuri and Rao 2004, 2013; Qian 2015; Wong and Guggenheim 2018), participatory approaches called community-driven development, or CDD, have become a central strategy for governments and international agencies over the past few decades. The World Bank alone supports more than 199 active participatory projects in 78 countries valued at USD $ 19.7 billion, representing a significant fraction of its lending portfolio (Mansuri and Rao 2004).

To shed light on this debate and to overcome some of the challenges in the existing literature, we create experimental variation along 3 dimensions. First, we examine whether or not there exists a collective action problem by comparing irrigation channel construction intervention with a cost-equivalent cash grant intervention. We randomize 192 of 240 villages to get the irrigation infrastructure and the remaining 48 villages to get untied funds transfer equal to the expected project cost, which we distribute equally to each constituent farmer within the command area. Second, we test whether community decision-making over the choice of the location of the channel matters via a one-farmer-one-vote ballot referendum initiative. We compare this with a design where the location is selected based on maximization of aggregate surplus using privately elicited willingness to pay from baseline, which we call the "data-driven benchmark". Third, we vary community agency over project implementation, where we either have the community constructing the selected channel on their own or third-party contractors hired by us execute the construction. We cross-randomize both choice and implementation dimensions to get a total of four experimental variations in the infrastructure program design with varying degree of community agency over choice and implementation, assigning 48 villages to each group. We stratify randomization by district to improve precision and statistical power as each district in our setting has distinct agro-climatic conditions and aquifer hydro-geology.

We note five main findings. First, we note the presence of collective action failure - a channel is constructed in 99% of the villages in the channel intervention groups (190 of 192 villages) whereas it is constructed in only 20% of the budget-neutral control villages (9 of 48 villages).[2] Farmers report substantially more days of total irrigation per acre relative to the control group. They receive an additional 40 days of irrigation per acre relative to a mean of 93 days per acre in the control group. The increase in access to irrigation is mainly driven by the constructed channel, which also directly abuts about 40% of all plots within the command area. There is also a modest, albeit noisy increase in groundwater irrigation, reflecting the recharge role of surface irrigation channel. The increase in the quantum of irrigation persists for 3 agricultural seasons (1.5 calendar years) following channel construction.

Second, we find that access to irrigation increases yield (output acre) and net revenue per acre with no substantial increase in the cost of production on average. However, the increase is only observed in data-driven location choice groups whereas there is no detectable difference between the referendum villages and the control group. Yield and revenue increases in the

[2] In 2 villages, some farmers objected to the construction at the last minute. We exited these villages.

data groups are 0.17 and 0.08 standard deviation units higher relative to the control group, where each farmer received INR 223 ($ 2.6) on average. The improvement in productivity from increased access to irrigation translates to an additional revenue equal to INR 2300 ($ 27) per farmer on average, implying a gain that is over 10 times the cash transfer.

Third, we record substantial differences in economic outcomes by choice groups but nothing significant when we vary how construction is implemented. We find that the channel selected via referendum is different from the aggregate surplus maximizing channel as in data-driven group in over half the referendum villages. This divergence could be plausibly due to relatively high costs of participation: few and a selective sample of farmers participate in the referendum. Surprisingly, some of those that participated in the referendum, voted differently from what they preferred in private, suggesting that the cost of voting alone cannot explain the observed differences. Consequently, referendum villages get 18 fewer days of surface irrigation relative to the data-driven benchmark (a 0.16 standard deviation decrease). This translates to productivity implications with referendum groups experiencing 0.2 standard deviation lower yields and 0.15 standard deviation lower net revenue per acre relative to the data groups.

Fourth, the referendum intervention generates regressive outcomes relative to the data-driven approach. Specifically, plots cultivated by large landholders experience more irrigation in referendum villages relative to data villages. These farmers reported almost no gains in additional days of irrigation at baseline but elicited higher value for certain field channels. Part of the higher valuation is mechanically driven by the fact that these farmers have larger plots and part plausibly driven by lowering of labor costs for paddy cultivation as suggested during informal interviews. Some farmers report that keeping the fields flooded reduces weed growth and subsequent weeding labor demand. We also observe lower costs of cultivation for this group, consistent with these anecdotes. The distribution patterns are also consistent with who votes - large landholders are more likely to vote.

Fifth, once a channel is selected, it does not really matter much how its construction is executed for subsequent water allocation or productivity implications. However, construction timeline is a little longer and we note minor deviations in the design of the channel when channels are constructed by the local communities (for example, community-led construction reaches the tail-end more often than in contractor-led implementation).

Viewing the results together suggests that having information on all individuals' private willingness to pay is sufficient to determine an efficient location for the channel with the objective of maximizing yields and net revenue. We infer that encouraging community participation is not a silver bullet and that the mode of giving agency to communities in the selection process matters. Further, giving budgetary support to individual constituents is insufficient to overcome coordination failure in creating and maintaining field irrigation channels. Importantly, the communities benefit from external facilitation for channel construction. We find that it does not really matter who gets involved in the actual construction process (i.e., whether local communities themselves or third-party contractors), at least over the time scale of this study. At the same time, external facilitation could induce some conflict.

We take a number of precautions in implementing the research design. First, we registered the entire pre-analysis plan before we executed any of the experimental interventions. Second, we implemented a strong data quality control protocol, where we went back to 25% of the respondents using different enumerator teams to collect some of the variables to test for any reporting errors. We did not find any substantial misreporting. Third, we triangulate our results based on survey reported data through random site visits and measurement based on visual inspections and GPS technology. Finally, we implement all our econometric specifications using difference in differences (DiD) estimator, measuring the effects as changes relative to baseline, which itself is balanced in expectation due to random assignment.

This paper makes several contributions. First, we provide experimental evidence on the presence of collective action and coordination frictions that leads to under-provision of commons. In our design, all treatment and control villages are fiscally similar, where each village receives funds equivalent to the cost of constructing a 300 meter-long field channel (INR 10000 or US$120). To our knowledge, this is the first experimental evidence that adds this existence result to a rich theoretical and quasi-experimental literature documenting collective action using observational and lab-experimental designs (Bardhan 2000; Montero 2022; Putterman 1981; Wade 1988). Importantly, we show that unconditional cash transfer, without facilitation, does not resolve the status quo failure of collective action. Furthermore, we show that conflict is one of the plausible causes for coordination failure.

Second, this paper sheds light on how best to design a CDD program. The typical design of an evaluation of CDD programs in the literature involves comparing a treatment group of villages that receives both facilitation and funding for an intervention to a control group that receives no intervention at all (Balán et al. 2022; Beath et al. 2017; Casey 2018; Casey et al. 2012; Khwaja 2004; Olken 2010; Qian 2015). Such design cannot, therefore, disentangle the impacts of the participatory approach via facilitation from the funding or the program itself and cannot address issues that are key to answering the research questions on the optimal role of local communities in development programs. While community-led approaches have been tried in contexts of conflict with weak or absent government where a comparison with more standard top-down implementation was infeasible (Beath et al. 2017; Casey et al. 2012), our context allows us to compare community-driven program design to one using data-driven approaches with third-party external contractor-led implementation. Furthermore, our evidence is based on apples-to-apples comparison of a specific set of infrastructure projects (field irrigation channels) using which we can ascertain whether community-driven program designs align or deviate from maximizing aggregate benefits. We also expand the insight by Khwaja (2004) (who finds that community participation in technical projects worsens project outcomes) in two ways. First, we hold the "technicality" of the project fixed in our experimental setting (which involves digging a mud trench as a field irrigation channel), thus accounting for "expertise". In our study, both the third-party contractor and local community member in charge of the implementation have similar "expertise". Second, we are also able to shed light on the interaction between the selection and implementation components of program design.

Third, we contribute to the literature on managing common pool resources (Rao and Shenoy 2023; Ryan and Sudarshan 2022; Sekhri 2011). Much of the literature has focused on groundwater extraction and use empirical strategies requiring strong assumptions on causal identification. We contribute by showing that even when there are no constraints in extracting water, as in the case of surface irrigation (Rodell et al., 2009), giving agency to local community in the choice of the channel location could generate regressive distribution. Using data-driven approaches to obtain private willingness to pay to identify an optimal strategy appears to be a better institutional design. With modern technologies for high frequency monitoring (Muralidharan, 2019), including remote sensing and detailed cadastral records with state agencies, estimating the marginal value of irrigation is plausible. This paper provides a proof of concept to show that there are instances when real-life implementation of local community choice, for example through ballot voting referendum initiatives, could diverge from data-driven approaches. In such a situation, policy makers can identify a more relevant approach that may be better suited to the context.

Finally, because participation may be costly, particularly for the poorest individuals, we study the optimal level of community engagement by including in the design two different degrees of community participation: the location decision that involves a referendum vote at a nearby ballot location and the actual construction of the channel that may require a few days of labor or an earth-moving equipment to construct the channel. We examine the time and money spent participating in each of these two activities, conflict that could arise within the command area (for example, over allocation of irrigation water or damages to plot boundaries during construction), as well as the (cumulative) benefits from each participatory component. Empirical evidence on the cost-effectiveness of community participation is limited, and to our knowledge, ours is the first study that tries to measure the cost of participation, including consequent conflict, separately for deciding the location from the cost involved in the actual construction of the project. This comparison isolates the role of community participation in choice relative to implementation (see e.g., Dal Bó et al. 2010 for a lab approach and Khwaja 2004 for descriptive evidence), where we find that there are differential costs of participation.

The rest of the paper is organized as follows. We describe the context of the study in greater detail in section 2 and elaborate on the research design and data sources in section 3. We document the reduced form results in section 4. We discuss the results in light of optimal institutional design and policy and conclude in section 6.

2 Context

We study surface irrigation allocation and provision of conveyance infrastructure through field irrigation channels (unlined, mud channels) connecting local village-level water reservoir (called minor irrigation tank) to downstream agricultural plots within its command area (see Panel A Figure 1 for a schematic). Tank irrigation is common in large parts of Southern and Western India, where surface run-off from rains are captured and stored in small-scale man-made reservoirs. These structures are complementary to groundwater irrigation through recharge from percolation. Thus, not only is surface irrigation cheaper socially (with lower cost of extraction compared to groundwater extraction that requires huge upfront investment and continuing electricity costs for pumping), it also complements groundwater resources for irrigation. Despite this, tanks and surface irrigation infrastructure are poorly maintained because they are a common pool resource. In contrast, groundwater is treated as a private good by those who are able to invest in their own private infrastructure (borwell and electric pumps) for water extraction even though it is also a common good as the underlying aquifer is shared between farmers or even villages, imposing significant externalities from water extraction. There are also currently no unified and comprehensive national or provincial policies on groundwater resource management in India as in the rest of the world.[3]

Collective action problem in the context of surface irrigation arises out of positive and negative externalities accruing to farmers based on the location of their plot within the command area and their access to alternate sources of irrigation such as groundwater. The problem is as follows: An upstream farmer (one with plot in the head-end) has a geographic advantage in accessing irrigation from the reservoir. Since paddy is tolerant to excess water but intolerant to its absence, head-end farmer ends up using more water than optimal. This reduces the amount of water available for downstream farmers, imposing a negative externality on them. Further, if a tail-end farmer invests in constructing a channel leading to his plot, he cannot exclude the head-end farmer from accessing water since the channel has to pass through head-end. This generates a positive externality on the head-end farmer whereas the cost is borne by the tail-end. This situation leads to a "race to bottom", leading to under-investment in the provision and maintenance of surface irrigation infrastructure.

In addition to concerns of status of surface irrigation infrastructure, there are no specific rules that are consistently followed in allocating irrigation water among farmers. In some villages, surface irrigation is continuous -"continuous flow irrigation"(Wade, 1988). In other villages, farmers allocate water using the depth of standing water on each plot. In Telangana, the consensus is that a farmer has the right to flood his or her plot until the standing water is 6 inches deep. This is known as "depth-based allocation". Other system such as "Warabandi" or "Taiband", common in north India, Pakistan, and few parts of southern India, involves a rotation system of irrigation allocation that is based on the number of days that the outlet to a

[3]Groundwater regulations was enacted as recently as 2014 in California and it's enforcement is currently being played out.

plot is open to the channel.

This collective action problem is exacerbated by the differences in the cost of access to irrigation via field channel, a function of distance downstream from the tank. All farmers in the command area grow paddy, and paddy farmers prefer to keep their fields flooded continuously, since paddy is intolerant to drying but highly tolerant to excess water. The uneven cost of access contributes to the unequal production outcomes as we find in a survey of about 2,000 farmers across 92 villages from a related study.

Table 1 presents the summary statistics across the study sample and by the different experimental groups. Each command area in a village in the study is about 100 acres of paddy cultivated areas, with 40-45 constituent farmers, each cultivating an average of 3.5 acres (mainly smallholders). There are 3 proposed location of channels within each of the command areas, each of which were deemed feasible to construct. The average valuation from constructing these channels amounts to around 30 additional days of irrigation, as elicited during baseline, amounting to a value addition of about INR 4500 ($55).

Our fieldwork and qualitative surveys shed light on cooperation failures among farmers. We find that constituent farmers rarely conduct public meetings, and even in a few contexts where farmers meet, it usually depends on local characteristics and history. In most villages we visited, the last meeting took place many years ago. Any maintenance work is usually allocated by an influential farmer (asami), or someone appointed by him.[4]. We also test this experimentally in the form of a budget-neutral control group, where we divide the average project cost - INR 10,000 - among farmers with plots in the command area equally via a private, in-person cash (INR 300-400 per farmer) transfer made to each farmer. Absence of external facilitation, high levels of internal cooperation would mean that the farmers would immediately pool their cash endowments into a “community pot”, where the funds in the community pot would be used to construct a channel at locations of community’s choosing.

[4]There is no clear pattern on who allocates volunteer labor, which has been observed in some villages while absent in many others

2.1 Conceptual Framework

A solution to overcome this problem is to internalize the externalities by a centralized authority, who would incur the cost of providing the common good as a social planner. However, the central authority can also leverage the utility derived from community engagement in the selection as well as construction process to improve the aggregate welfare. This ancilliary welfare improvement could arise from monitoring the quality of construction, volunteer labor contribution, and ongoing maintenance from the local community to ensure that the good lasts multiple seasons.

Two key levers to induce community participation in any development program are giving communities a choice over the location of the infrastructure project, or giving them agency over implementation or a combination of both, which helps overcome two important challenges. First, giving community agency over choice of the good (i.e. location of the surface irrigation channel) resolves the problem of information asymmetry - that of eliciting local preferences, which may not be observable to a social planner. A common method of eliciting local choice is through a referendum, which provides equal weights to the preferences of all farmers irrespective of their baseline value from increased irrigation.

On the other hand, the social planner could collect data on private willingness to pay to address this concern, especially due to the imperfect weighting and other strategic behavior in the referendum process.

Second, involving the community in construction of the public good accounts for the agency problem present among the agents of the central authority, i.e. the construction workers or the project manager. Community involvement ensures that the construction is of good quality (for example, uniform width and depth of the channel) and voluntary contribution of labor or other inputs that maximizes the aggregate surplus. Both these aspects could increase overall community satisfaction with the public good.

We test for these alternate theories of institutional design to resolve the collective action problem in the context of irrigation infrastructure maintenance and allocation in this paper.

3 Research Design and Data

We conducted a randomized control trial (RCT) in which we varied (a) whether the location of the channel was chosen by the community through the referendum via ballot voting or based on data by a central agency (i.e., the research team) and (b) whether the construction of the channel was undertaken by the community or a third party contractor hired by the central agency (i.e., the research team).

3.1 Description of treatments

The intervention followed a 2-factorial stratified RCT varying how the location of the channel is chosen (choice-arm), crossed with how it is constructed (implementation-arm). All treatments were randomized at the tank-command area-level and since we only included one tank in the study per village, treatments are also randomized at the village level. We stratified the randomization by district since our sample of villages spanned multiple districts with different agronomic conditions. The experimental groups arising from this include:

1. T1 (Referendum Location Choice and Contractor-Led Implementation): Location chosen by local village community via a referendum (simple majority) with channel construction carried out by an external contractor identified by a central agency.

2. T2 (Referendum Location Choice and Local Community-Led Implementation): Location chosen by local village community via a referendum (simple majority) with channel construction also carried out by local village community.

3. T3 (Data-Driven Location Choice and Contractor-Led Implementation): Location chosen by a central agency with channel construction carried out by an external contractor identified by the same agency.

4. T4 (Data-Driven Location Choice and Local Community-Led Implementation): Location chosen by a central agency with channel construction carried out by local village community.

In addition to the 2x2 matrix implied by the two-part randomization, the design also includes a budget-neutral control group (C). Farmers in this control group receive the cash equivalent of the per-capita cost of the total project but the choice of location or construction are not facilitated.

3.2 Sampling and Baseline Data Collection

Based on extensive piloting and using secondary data from the same context, we calculated that two hundred villages with 30 farmers/plots per (village-level) command area generated statistical power at 80 percent under significance level at 5 percent to observe a modest treatment effect using access to surface irrigation as a key outcome. We simulated statistical power using the estimated empirical effect sizes from pilots we conducted in similar regions.[5] We included more villages than the required sample size to allow us to account for any compliance or take-up issues at the time of intervention. We sampled 20% more villages/tanks than required, generating a study sample of 240 villages.

We recruited villages with tanks that had a feasible command area (around 100 acres) for surface irrigation to be viable. Furthermore, our population of interest are tanks where surface irrigation was feasible, was done in the past but where channels were currently under state of disrepair. Thus, our first step was village recruitment. We used the irrigation tank census data from publicly available data source provided by the Government of Telangana. We focused on tanks with at least 100 acres of command area and where sluice gates were operational.[6] We sent field enumerator teams to villages that satisfied these criterion in order to physically assess feasibility of surface irrigation channel construction via random audits. We also secured community-level consent from key stakeholders in the villages to participate in our study so that we could recruit the required sample size of 240 villages where surface irrigation channels were feasible to construct and formed the sample of this study.

At the start of the study, we carried out schematic mapping of the irrigation tank and its command area with a group of farmers knowledgeable about the topography in all 240 study villages. Additionally, we also listed the census of agricultural plots within the command area

[5] We note that during the pilot we had 100% compliance. All treated villages constructed the channel while none of the control villages did.

[6] A non-functional sluice gate would require more technical civil engineering-based intervention that is typically only provided by the state government.

and identified the cultivators for each of the plot using official documents from the local village council. This listing census data forms our sampling frame from which we randomly sampled plots (and its cultivators), stratified by relative geography within the command area (i.e., whether in the head-end location or tail).

An important aspect of the mapping exercise was to identify 2-3 feasible locations between the source of the tank (sluice gate) and the command area where we could construct surface irrigation channels during the mapping process. For each village, we created maps (as in Panel B Figure 1) depicting the downstream command area, sluice gates along the tank, and feasible locations for constructing the surface irrigation channel. These maps were very important aspect of implementing our research design and data collection. For each feasible location, we collected data on the approximate length of the channel and expected costs of construction if the village communities had to hire labor or an earthmover to construct a channel at the location. Due to budget limitations, we capped the project specification, i.e. the specification of construction at any given location, to be within INR 10,000. In practice, this budget limit was binding in all 240 study villages.

Using the plot census data, we drew a stratified random sample of 30 plots per command area, stratified by plot location (i.e. head, middle, and tail). We used stratified sampling in order to generate a representative sample of all plots within a command area. This is because we observed substantial differences in economic outcomes and bargaining options faced by cultivators based on the location of their plot. We attempted to draw equal number of plots from head, middle, and tail-end locations of the command area. This generated a sample of ≈ 6200 sample plots (and farmer cultivators) across 240 study villages.[7] The sample of plots largely map one-to-one to sample of farmer cultivators within in the village.[8]

We carried out a baseline survey of the sample respondents - farmer cultivators of the sampled plots, recording details of their cultivation decisions and outcomes for the preceding two agricultural seasons along with demographic information of the farmer household and the study village. In addition to the standard agricultural cultivation and demographic questions, we gathered data on baseline level of cooperation and conflict over irrigation resources and

[7]In villages where the entire command area population of plots was less than 30, which happened in a small fraction of villages, we surveyed the entire command area.

[8]Only a handful of farmers cultivated multiple plots within the command area. Since we sample at the plot-level, we interviewed the same farmer if more than one of their plots were sampled.

baseline level of trust between farmers within the command area. Lastly and importantly, we elicited each of the sample farmer's private "willingness to pay" or value accorded to each of the feasible locations to construct the channel. Specifically, we noted which of the locations would be closest to the sampled plot, how many additional days of surface irrigation would they expect if a channel at each of the feasible locations were built, how much additional revenue would they earn, and finally what would be the total value in terms of net revenue after accounting for costs. The details of the script used to elicit this information is available in subsection A.2.

Plot-level Sample Expansion: Post-construction of the irrigation channel, we additionally surveyed all plots along the constructed irrigation channel not covered by our baseline sample of plots, including along the channel that could have been constructed (as per elicitation data and thus serve as counterfactual) in the control group.

3.3 Selecting Channel Location

Following baseline data collection exercise, we calculated aggregate valuation for each location for the channel across the representative sample of farmers using their privately elicited value for each location. Thus, for all 240 study villages, we identified which location would maximize aggregate valuation across all plots within the command area. In the "Data-Driven Location Choice" groups (T3 and T4), we identified this location to construct the surface irrigation channel.

Next we conducted referendum exercise as a secret ballot one-farmer-one-vote in the "Referendum Location Choice" groups (T1 and T2). In villages assigned to these groups, we gave each farmer cultivator a ballot card with the location of the 2-3 feasible channels identified earlier on, covering all those farmers associated with the universe of plots within the command area (see Figure A3 for an example of the front and back of the ballots we handed out). We told these farmers to select one location of their choice (one farmer-one vote) by indicating their choice on the card and to deposit the card in a locked ballot box kept at a prominent place before the end of 3 days from the time of our communication.[9] We also gave them copies of the command area schematic map after explaining the map again and identifying the location of their plot on

[9]Due to COVID19 pandemic, we could not conduct a polling day in the village that would have led to queuing up of farmers to cast their votes.

the map. We ensured that the voters understood the purpose of the referendum and that their votes would be counted to identify the location that receives majority votes, which would then be constructed.

It is plausible that the aggregate value maximizing channel location would coincide with the referendum winner if everyone showed up to vote (or at least a representative sample of the entire voter population voted) and voted according to their private elicitation. It is also plausible that the channel location selected under the two experiment groups - referendum vs. data-driven location choice groups - could be very different and could result in different economic outcomes. For example, if those who show up to vote are a highly selective sample or if farmers don't necessarily vote for channel that map to their baseline private valuation. We examine whether these hold empirically - that is, who votes and how they vote.

3.4 Constructing the Surface Irrigation Channel

Once the location was selected either by vote or by identifying the aggregate value maximizing channel, we communicated the outcome publicly in the village. Other than by loud public announcement, we also put up a poster at a prominent public location that indicated the selected location and how it was selected (whether through majority votes or by us). In the same announcement, we indicated whether the channel at the said location would be constructed by a third-party contractor (that we would hire) or if we would provide funds to a member from the community to construct the channel themselves. We asked the community to suggest a member to serve as our liaison based on consensus.

In "Contractor-Led Implementation" groups (T1 and T3), we hired an external contractor with experience in digging field irrigation channels from the sub-district-level market place. We ensured that this external contractor was not from the village and to the best of our knowledge, had no social or economic links with the village. We sought approval from the village council head and other key leaders to begin the construction on a said date [10] before starting the construction. The construction itself was completed within a day or at max two days. The research team members were present at the site of construction to manage the third party contractor.

[10] As per local traditions, all groundbreaking activities are expected to started at an "auspicious" time with a formal groundbreaking ceremony.

In "Local Community-Led implementation" groups (T2 and T4), we appointed a local community member who volunteered to be our liaison and who had experience overseeing construction projects. We informed the community who this liaison was and transferred half the average project cost to the liaison's bank account at the start of the construction and the remaining part based on the actual costs borne at the end of the construction after an independent audit. Beyond providing funds and follow up reminders to complete the construction process, neither the research team nor any external third party did any further facilitation of the construction process. In order to ensure that the liaison didn't evade or failed to honor the construction contract, we obtained their ID, phone number, and bank account details. In practice, we didn't face any contract violation in any of the villages, plausibly because of their reputation concerns among the community of farmers with plots in the tank command area.

3.5 Implementing Budget-Neutral Control

In the control villages (group C), we distributed cash equal to the average expected cost of the project across all 240 villages ($\approx$ INR 10,000) equally among all farmers with plots within the command area of the irrigation tank. We did this around the same time as channel construction intervention to ensure that the timing of any intervention was similar across all experimental groups. While distributing the cash, we again mentioned our study's objective about understanding surface irrigation. We highlighted that we were equally distributing cash across all farmers with plots in the command area and emphasized that each farmer could decide how best to use the cash provided.

3.6 Empirical Specifications

We estimate the effects using pooled specifications where we pool the specific choice and implementation arms - for e.g., just examining the effect of referendum location choice and data-driven location choice relative to the budget neutral control group, or examining the effect of local community-led implementation and third party contractor-led implementation relative to the control. The idea behind pooling is to enable specific hypothesis tests examine the role of community involvement in choice and their role in implementation separately before examining the interactions.

$$Y_{ivdt} = \delta_d + \delta_t + \phi_1 \text{Referendum Loc}_{ivdt} + \phi_1 \text{Data-Driven Loc}_{ivdt} + \epsilon_{ivdt}$$
$$Y_{ivdt} = \delta_d + \delta_t + \gamma_1 \text{Local Community Imp}_{ivdt} + \gamma_2 \text{Contractor Imp}_{ivdt} + \varepsilon_{ivdt} \quad (1)$$

We also estimate the empirical specification in long-form as follows:

$$Y_{ivdt} = \delta_d + \delta_t + \sum_{j=1}^{4} \beta_j T(vt)_j + \epsilon_{ivdt} \quad (2)$$
$$Y_{ivdt} = \delta_d + \delta_t + \sum_{j=1}^{4} \beta_j T(vt)_j + \mathbf{X}_{ivdt}\Gamma + \epsilon_{ivdt}$$

In this equation, d indexes the district, v indexes the command area (a village) in our sample, i indexes sample farmer (some specifications are at the command area-level itself), and t indexes data collection round/agricultural season. The leave out group is the budget-neutral control group. All the treated variables switch on after the baseline round. Standard errors will be clustered at the level of treatment assignment, i.e. at the command area-level (Abadie et al., 2022). Our base specification does not include any controls since the specification is implemented as a Differences in Difference (DiD) design. For robustness, we will include baseline variables that remain unbalanced as controls, where $\mathbf{X}_{ivdt}$ represents a vector of such baseline characteristics.

Predetermined Variables as Controls We use the following variables for balance check after random assignment: tank area; tank storage capacity; command area area; history of past repairs; number of plots in command area; number of cultivating farmers in command area; number of feasible locations of irrigation channels within command area; value of highest elicited channel; farmer (cultivator)-level demographics - age, gender, jati, total landholding, total irrigated land, sample plot-area, presence of borewell on plot.

In all our empirical specifications, our preferred specification is one without controls. This enables us to interpret the coefficients as average treatment effect or ATE.[11] We carry out multiple robustness tests to ensure that the estimated coefficients do not suffer from any biases,

[11] Randomization ensures that the ATT parameter estimated using the DiD estimator coincides with ATE.

particularly if some variables are individually unbalanced at baseline. First, we account for the measures that remain unbalanced even after random assignment of experimental status. Additionally, we control the following in robustness as pre-specified: (a) deviation between the channel selected from the ballot exercise match the highest-valued channel from the private elicitation exercise at the aggregate command area level, (b) demographic details of persons in-charge of channel construction under bottom-up construction, (c) a summary index for baseline measures of collective action - joint sale, joint input purchase, joint investments (e.g. transportation to markets), and baseline-level of trust between farmers.

Second, in order to discipline our selection of control variables, we implement the post-double-selection method of (Belloni et al., 2014) to identify the subset of the control variables from above.

We account for multiple outcomes using two approaches: First, we generate a single summary measure (index) for each class of outcome in standardized units, either using Principal Component Analysis (Kling et al., 2007) or standardizing the variables in the index and then adding them. Second, we test each family of measures (under different groups presented under the data section - command area-level measures, farmer-level measures, and plot-level measures) jointly using multiple hypothesis corrections using FDR q-values.

3.7 Data

We collected data using primary surveys administered to farmer-cultivator of the sample plots, and investigator-led random audits with our field implementation partner J-PAL South Asia at IFMR. Specifically, we collected the following self-reported (survey) data: (a) a set of feasible locations for channel construction within the command area from key informants, (b) a listing census of command area plots and its owners, (c) a baseline survey containing private elicitation of preferences over the feasible set of channel locations under (a), (d) the number of valid ballots cast for each channel in ballot villages, (e) follow-up survey data collected via phone surveys on various satisfaction and cost of participation measures during and post intervention, and (f) a comprehensive endline survey containing detailed agricultural production measures for all cultivation seasons during the year following the intervention, including wet (Kharif) as well as dry (Rabi). Finally, measured the quality of channel construction including the dimensions of constructed channel relative to the initial specification through random site audits and measurements by trained enumerators.

A key innovation of this study is the measurement of baseline preferences (stated willingess to pay or WTP) over a list of all proposed channels by our sample farmers that are feasible for implementation. Specifically, farmers in the command area were asked to provide a valuation for each channel based on the number of additional days of surface irrigation and additional revenue earned. From this, we identify the channel that maximizes aggregate surplus within an command area. This is unique in the literature where most studies on community participation do not document ex-ante preferences of each community member. In the absence of such a system, selection of public goods has hitherto been determined in an ad-hoc manner by either village elites or through Gram Panchayat (elected village council), which may only cater to the preferences of select groups at the expense of collective surplus.

We collected several primary and secondary outcomes as well as intermediate outcomes related to the construction of the channels as specified in our pre-analysis plan detailed in Appendix B. We examine the effect of the program design on key outcomes at all stages of the causal chain, starting with the first stage, construction and community participation related measures. In terms of economic outcomes, we examine access to irrigation (in total and separated by source), agricultural productivity (yield per acre), and net revenue. Since over 90% of the farmers within the command area cultivated only one plot of land (that within the command area), these outcomes also reflect outcomes at the household-level. For a small share of farmers that cultivated land both inside and outside of the study command area, we measured aggregated revenue and cost implications at both sample plot and at farmstead levels.

4 Results

We discuss the Intent-to-Treat (ITT) estimates of the channel intervention program including the additional experimental variations in the program design along the choice and implementation margins of community participation. We present results along the entire causal chain, starting with the selection of channel location, the dynamics of program implementation, and finally concluding with economic (production-related) outcomes.

4.1 Baseline Summary Statistics

Table 1 presents the summary statistics for the study villages including the size of village-level cultivated and irrigated areas, number of farmers cultivating within the command area, the number of feasible locations identified for constructing a channel (including total number of channels and percent in disrepair), area cultivated by representative farmer, additional days of irrigation and additional revenue implications from surface irrigation channels, and stated valuation (WTP) for the different feasible locations. Cultivated and irrigated areas are large in our study villages with over 80% of the cultivated lands that are irrigated. There are 40-45 farmers cultivating an average 3.5 acres of land within the specific tank command area included in this study. The villages proposed 3 feasible locations for constructing a channel on average, which would bring 30 additional days of irrigation (in expectation at baseline). The average value for a location was under INR 5000 per farmer, roughly mapping to the revenue implications. These variables are fairly balanced across the experimental groups and none of these variables have strong predictive power in determining the specific experimental group assigned to a village.[12]

Figure 2 shows farmers' stated willingness to pay strongly correlate with the irrigation and revenue implications from their most preferred channel. We also examine irrigation, revenue, and stated WTP by baseline farmer and plot characteristics observing that the channels increased access to irrigation for smaller farmers (with < 5 acres of land), those without borewell irrigation source, and those from marginalized caste groups. Revenue per acre implications are also similar, suggesting that the maximum revenue implication of increased surface irrigation are most likely felt by farmers with limited access at baseline - small farmers from marginalized groups without borewell cultivating in the tailend region within the command area. The stated WTP for the channel incorporates total net revenue effects after accounting for costs. The correlation between WTP, which is the total surplus at the farm-level, is positive with respect to baseline farm size, some of which is mechanical due to larger land size and some could be due to potential cost savings if irrigation and production costs are complementary. On the other hand, this correlation is either not significant or is negative for headend and female farmer cultivators.

[12] The joint orthogonality test of balance is conservative by over-rejecting the model (Kerwin et al. 2024) in comparison to pair-wise orthogonality tests.

We also measured the extent of water trade, coordination, and conflict between farmers over irrigation at baseline. We document the summary statistics in Table 2. Under 60% farmers report trading water for irrigation, and around 50% state that they coordinate with others in the village and command area around irrigation and agricultural production. 7-9% farmers have experienced conflict with their neighbors over irrigation but under 20% villages had any rules over sharing water. Even though only a few farmers report conflict (plausibly under reporting due to biases), the fact that only a small fraction of villages have either formal or informal rules over water sharing suggests coordination failure.

4.2 Characteristics of Channel Selected

In data-driven villages, the aggregate surplus maximizing channel was selected by default. This channel abuts nearly 50% of the command area farmers but is not significantly correlated with any of the demographic characteristics (Figure 2). We announced the selected location publicly in villages assigned to data-driven groups once referendum process had concluded in the referendum villages so that the timing of announcements was similar across these experimental groups.

The channel selected through referendum was influenced by who voted during the referendum and how they voted (see Table 3, Table 4, and Figure A4). Looking at the characteristics of farmers and sample plots along channels ordered by their aggregate value or by the number of votes, we do not find any specific baseline characteristics to be correlated (Col 1-3 Table 3). However, these characteristics are correlated with turnout, which was around 43%.[13] Specifically, we find that those that would have benefited from a channel did not vote (Col 4 Table 3). For example, we see headend farmers with higher turnout despite lower expected irrigation effects or lower stated WTP for any channel at baseline. We also see large farmers more likely to vote who also reported higher stated willingness to pay at baseline.

Overall turnout is surprisingly not explained by the share of farmers along the aggregate surplus maximizing channel but rather another channel, which plausibly has the second highest surplus. Though the channel that gets maximum votes is a consequence of voting, we note

[13]The turnout number is in the generally observed range in local and national elections in India. See the Election Commission of India website (`https://elections24.eci.gov.in/docs/BnS4hhbvK9.pdf`) for turnout statistics in national and state-level elections.

that farmers in a specific geography within the command area coordinate to vote for the channel that is closest to them (see Col 1-2 Table 4).

As a consequence of this differential voting behavior, the location selected via the referendum was not the aggregate surplus maximizing location in over 50% of the referendum villages. However, in some villages where there is a greater share of farmers along the channel receiving highest votes, we notice that the referendum outcomes are more likely to align with the data-driven channel (see Col 4-6 Table 4).[14]

Panel A Figure 3 shows that the referendum process was well understood and that farmers mostly selected channels based on what would be best for them at that point in time. However, a substantial share of farmers (close to a third) indicated that they voted for channels that was beneficial to other farmers or based on changes induced by weather or other construction events within the command area. This likely explains why many farmers who voted ultimately did not vote for the channel that maximized their own private WTP elicited during baseline.

As a precursor to our intervention, we had posed a hypothetical question to the farmers asking how should one choose the ideal location for constructing the channel during baseline. Panel B Figure 3 shows the distribution of these responses. The majority - over 45% of the sample farmers - suggest that the location selected should be the best for the *most* farmers. A substantial others - 40% - said that they would let the elite farmers decide which location should be selected. Surprisingly, almost none suggested that selection should be done via referendum voting. This is consistent with farmers' response on whether they found the selection process to be free, fair, or transparent after the location was selected and announced publicly. While a large share of farmers (over 90%) in both referendum and data-driven groups said that the selection process exhibited these qualities, fewer farmers in the referendum groups agreed with the statement relative to farmers in the data-driven groups. This difference is statistically significant at 1% (p-value < 0.01).

Taken together, these results on channel selection suggest that participatory methods like referendum voting may not best capture selection of channel location that generates highest aggregate surplus for farmers within the command area. We describe the production implications of this location choice later in this section.

[14] A 1% increase in the share of farmers along the channel with the highest votes corresponds to a 1% increase in the likelihood that the aggregate surplus maximizing channel, as in the data-driven groups, is selected.

4.3 Characteristics of Channel Constructed

99% of the channel intervention villages constructed the selected channel (either the one selected by the referendum or one by data) compared to only 20% or 9 (of 48) control villages, where they constructed some channel in the command area but none that coincided with the aggregate surplus maximizing channel. The two villages (of 192 treatment villages) where we did not construct include one from referendum local community implementation and another from data-driven contractor-led implementation groups each. In the referendum village, some of the farmers did not agree with the referendum outcome and blocked construction. In the data-driven village, a few village elites prevented us from constructing the selected channel. Following the consent procedure, we exited these villages even though the farmers and communities themselves had consented to possible construction at baseline and had participated in the referendum process. This attrition, however, is not correlated with the treatment status as noted in Col 1 Table 6.

The cost of the program was relatively similar across the treatment and control villages at baseline (estimated costs are no different between either of the experimental groups as seen in Panel A Figure 4), although we experienced minor cost overruns during the construction process. This was mainly due to increased rental and wage costs of inputs (capital and labor) involved in the construction. In control villages, we used the estimated cost of constructing the aggregate surplus maximizing channel and divided it equally among all farmers within the command area. We had to round-off the cash value to the nearest INR 50 given the logistical challenges of cash distribution as cash denominations were only available in multiples of 50.[15] As a result, the actual cost per capita is slightly but not meaningfully lower in control villages - a difference of INR 30 per farmer on average or a village-level difference of $\approx$ INR 1500 (see Panel B Figure 4). The cost overruns are similar across all channel intervention groups and none of the location choice and implementation groups differ significantly in the per capita realized costs.[16]

[15]Many farmers did not have digital bank accounts at the time of our intervention and as a result, the cash transfers were in person, physical transfer of currency notes. For example, if the computed cost per capita was INR 220, we provided INR 200 in cash to the farmer after rounding to the nearest 50. Similarly, if the cost per capita was INR 280, the rounding resulted in INR 300 being paid to the farmer.

[16]For robustness, we control for cash per capita in our regressions of economic and production outcomes and note that the results continue to hold even after accounting for these differences.

Figure 5 presents results around the construction process itself, including which channel got constructed and how. In 93% of villages in data-driven groups, the aggregate surplus maximizing channel is constructed whereas this channel is constructed only in 45% of referendum villages. This difference is stark and statistically significant (Panel B Figure 5). On the other hand, similar proportion of villages (≈ 70%) assigned to the different implementation groups construct the surplus maximizing channel and the difference between these groups is a precise zero (Panel B). Local community-led implementation strictly follows the selection of location and does not deviate from the exact channel to be constructed.

Quality of construction, measured as an index incorporating the length of the channel, uniformity of the channel dimensions, whether reaches until tailend, and whether we noticed any repairs and modifications, is better in control and local community-led implementation groups. Local communities are better at constructing channels themselves compared to the third party contractor, and the difference in the quality of construction between these groups is significant at 10%.

Lastly, we noticed differences in the mode of construction between the location choice and implementation groups (Panel C Figure 5). In both referendum and local community-led implementation groups, the key community member in-charge of the construction is more likely to be related to the elected representative of the village council (indicating elite status) and is less likely to own a construction machine themselves relative to the corresponding counterfactual. On the other hand, they are no more or less likely to carry out construction using capital or labor. Additional construction-monitoring data (the number of visits, difficulty in completing the intervention, timeline delays, and arguments with the field research team) indicates that the construction was relatively easier to implement in the referendum groups whereas it took more time to complete construction in the local community-led implementation groups (Table 5).

4.4 Access to Irrigation

Figure 6 and Figure 7 depict the effect of the different channel interventions relative to the cost-neutral control on the total number of days of irrigation per acre as well as irrigation days disaggregated by source (surface vs. groundwater). We note a substantial increase in the total number of irrigation per acre days and number of surface irrigation days per acre (over 30 additional days) among both dimensions of program design in channel construction groups compared to the cost-neutral control group. The increase in total irrigation is mainly driven by increase in surface irrigation although we also notice modest increases in groundwater irrigation per acre days as well (around 10 additional days but not statistically significant). The increase in the total number of days of irrigation is similar in ballpark with the expected increase elicited during baseline, minimizing concerns of response bias.

Second, the increase in irrigation is lower in referendum groups relative to data-driven groups where aggregate surplus maximizing channel was built. The magnitude of difference is around 18 fewer days or a 0.2 SD units in effect size relative to the data-driven groups.

Third, we find no difference in irrigation outcomes based on who constructs the channel. The increase in irrigation days per acre is similar across both implementation groups. We are unable to reject any differences lower than a 0.02 SD effect size, which is small enough to infer that there are practically no differences in irrigation outcomes due to the actual implementation process.

4.5 Agriculture Production

We record substantial differences in production outcomes across the experimental groups. Figure 8 shows that output per acre (yield) increased in the data-driven groups relative to control. This is translates to a 0.17 standard deviation unit improvement in yield relative to status quo. In contrast, referendum groups experience no average improvement in yield relative to control and a correspondingly lower yield advantage compared to data-driven groups. The difference in yield between location choice groups is statistically significant, where referendum groups experience significantly lower yields than data-driven groups. The difference in yields is not significantly different across implementation groups.

The effect on net revenue is noisy when compared to control group although the magnitudes indicate a plausible positive effect of data-driven location choice and a negative effect of referendum-based location choice. The difference between referendum and data-driven groups are significant, suggesting an average revenue lowering effect of the referendum intervention. The revenue effects of implementation groups are noisy although there is suggestive positive effect of local community-led implication. Given the absence of irrigation and yield effects, the increase in revenue could plausibly due to side transfers from the funds provided to the local community for construction and not real productivity effects from increased irrigation.

On cost of production, there is no substantial difference between any of the experimental groups relative to control on average. Paddy is historically grown in irrigated areas such as river deltas or within the command areas of irrigation schemes such as in our context. Local agricultural experts in our context refer to paddy as a "lazy farmers' crop" because of its high water tolerance relative to weeds. As a result, farmers prefer keeping their paddy fields flooded if they have access to irrigation to reduce the need for weeding labor. This suggests that there may be large heterogeneity in the effects on cost of production by baseline characteristics even though the average effects on total and labor costs are modest and noisy.[17] We discuss this more when examining the distributional implication later in this section.

4.6 Status of the Channel at Endline

We went back to the study villages three agricultural seasons (1.5 calendar years) following the construction of the channels to examine the status and use of the constructed channels. We carried out unannounced, in person audits of the tank command area and constructed channels, speaking with whoever were present at the site. Table 6 documents the results from this exercise. We were able to carry out our inspections in 238 of 240 study villages (2 were dropped due to withdrawal at the time of construction), with no differential attrition by experimental status. Overall, more than 60% of the constructed channels in the treated villages continue to be functional during these visits in terms of length and its ability to transport water all the way to tailend regions.

4.7 Direct Costs of Community Participation

We asked questions about cost of participation, both in terms of time-use, including voluntary contribution of labor, and any additional cash matched-up by the constituent farmers. We find that the costs are much lower in all of the treatment groups relative to the cost-neutral control group, where about 20% of the control villages constructed field irrigation channels (see Fig-

[17] At the village-level, we find suggestive evidence on lower agricultural wages to female labor who are mainly engaged in weeding labor (Table A1).

ure A5). We surmise that external facilitation could have reduced the coordination and communication costs required for collective action among constituent farmers in channel intervention villages whereas in control villages, the farmers would have had to mobilize and collectivize to construct the channel they decided to build. This is seen more in terms of time cost rather than direct monetary contribution to the construction of the channel. There are no differences in these cost either by choice groups or implementation groups.

4.8 Distributional Implications

In addition to the average effects of the interventions, we find distributional implications of referendum program design. We observe heterogeneity in irrigation allocation as well as in production outcomes by landholding size and other preregistered baseline characteristics. Figure 9 presents heterogeneous treatment effects on irrigation and production outcomes by baseline landholding size categorization as large landholder. We note noisy albeit higher irrigation per acre days and higher net revenue per acre and on the other hand, lower costs of production per acre among large landholders relative to their counterparts in data-driven groups. Heterogeneity by other other characteristics are unclear - plots with borewells receive more surface irrigation compared to plots without borewells but these don't translate either to better yield/revenue or lower costs of production. We find no heterogeneity by plot location - head-end plots are no different from plots elsewhere in the command area with respect to irrigation or production outcomes.

Interpreting these results with the outcomes from the referendum process, we surmise that it is plausible that the differential selection of channel location could be advantageous to certain groups of farmers, who experience higher irrigation and/or other production outcomes relative to the average farmer. While we need additional evidence to conclude this as an implication of elite capture, our results at least caution us from inferring positive implications from participatory processes such as referendum voting.

5 Economic Mechanisms

What explains our main reduced-form results? Compared to a private cash-grant of INR 223 in control, the surface irrigation intervention led to an increased agricultural income (net revenue) of around INR 2400 per farmer in the command area. Why were surface irrigation channels not already provided given the real economic benefits from it?

5.1 Conflict as Coordination Failure

Anecdotally, farmers in many villages reported at baseline that they kept the sluice gates shut and avoided surface irrigation altogether due to conflict between farmers in the command area. As a result, tank was only used for recharging groundwater aquifers, which in reality only benefits those with access to groundwater irrigation sources like borewells. Borewells are expensive to invest in and is a risky investment. There is a substantial probability that a drilling for borewell yields no water and/or existing borewells can eventually go dry. Thus, the existing equilibrium is not only a Pareto inferior equilibrium but also an unequal one.

At baseline, one of the reasons for lack of surface irrigation is that over 45% of channels are under severe disrepair. Despite this, only 15% of the villages have any formal or informal rules over irrigation water sharing. 7-9% of the sample farmers reported experiencing conflict with their neighbors over access to irrigation and only 50% of the sample farmers report engaging in some form of coordination while the other 50% do not coordinate. While these numbers could plausibly be understated/overstated respectively, it suggests that there are multiple frictions to ensure successful and sustained coordination over managing local irrigation resources (we discussed these statistics in detail in subsection 4.1).

5.2 Conflict Reporting in Surveys

Overall, we notice higher albeit noisy increase in reported conflict in any of the channel intervention groups relative to the budget-neutral control. Specifically, reported conflict is higher when the implementation is led by a third party contractor whereas the mode of selecting the location for the channel does not affect conflict. Further examining the interactions between location choice and construction process, data-driven location choice group with third party contractor led implementation has highest reported conflict although given reporting bias, the estimates are noisy to be compared between the different experimental groups. We try to minimize the noise by controlling for baseline reported conflict in an ANCOVA specification but the results continue to be noisy (see Figure A7). We corroborate this result on reported conflict with results from non-incentivized trust and dictator games that we played during both baseline and endline surveys.

5.3 Evidence from Trust and Dictator Games

While stated conflict measures could have reporting bias, we measure the extent of trust and pro-sociality among farmers within the command area using non-incentivized trust and dicatator games. In the trust game, we provided the following vignette: "Imagine the following hypothetical situation. We offered Rs. 100 to some farmer in the head/tail location in this command area, allowing them to share any amount from this Rs.100 with you. Any amount they share with you, we double the amount from our side before giving it to you. Your partner decided to share Rs. 50 out of Rs. 100 with you. We doubled this from our side and are now offering you Rs. 100. Now you may choose to keep all the money and the game ends. Alternatively, you may choose to send any amount out of Rs. 100 you received back to your partner. We will not be doubling this amount and so, the amount you choose to send to your partner is the amount they will receive." The objective of this exercise was to measure reciprocity of trust. We set up the trustor (player 1, who is framed as an actual farmer within the head or tail regions in command area) to be neutral in terms of how trusting they are. In the vignette, they share half their endowment, suggesting that they were neither fully trusting nor showed a lack of trust. The action of the sample respondent thus captures the action of player 2 (trustee) in the trust game. If they returned a higher amount, it would mean higher reciprocity. On the other hand, if they returned nothing, it should indicate lack of reciprocity.

In the dictator game, we provided the following vignette: " Now consider another hypothetical situation. Suppose we offer you Rs. 100. You can share a part of this amount with any farmer in head/tail location in the ayacut. They can either accept or reject this offer. If they accept, they will receive the amount you share and you keep the remaining. If they reject the offer, both of you will receive nothing. Suppose you decide to send Rs. 20 to this person and they accept. In this case, they will receive Rs. 20 and you get to keep remaining Rs. 80. If they reject, both of you will receive Rs. 0. Of course, you can choose to send any amount. Rs. 20 is just an example." The objective of this exercise was to assess the extent of pro-sociality among the sample farmers towards a representative farmer in the head or tailend regions of the command area. A higher amount of transfer would indicate higher levels of pro-sociality towards the head/tail farmer.

Albeit not incentivized, we account for baseline levels of stated contributions with endline contributions and attribute the residual effects to coordination frictions. As with reported conflict, we carry out ANCOVA analyses, where we regress contributions in trust and dictator games by sample farmer to a real farmer with plot either in the head or tail regions within the command area. In both the games, we asked for contributions to these two types of farmers to each sample respondent.

Figure 10 presents the differences in contributions across the sample farmers by experimental groups and compares within choice and implementation groups. We note three main patterns: First, contributions in both games are lower in all channel construction groups relative to the budget-neutral control to both headend and tailend farmers. Lower contribution in the trust game suggests that the sample farmers are less trusting of other farmers in the command area. Lower contribution in the dictator game suggests that the sample farmer are less altruistic towards other farmers in the command area. Second, gap between contribution to a head or a tail end farmer is similar between the referendum and the control groups in both trust and dictator games. However, the gap is lower in data-driven groups compared to the control group (and significant at 5% for trust game). This result on the contributions gap is driven by lower contributions to the head-end farmer relative to tail-end farmers in both trust and dictator games in data-driven groups. Third, the contributions are meaningfully different (0.1 SD) between the choice groups (referendum relative to data-driven location choice groups) whereas they are similar across construction implementation groups (local community-driven relative to contractor-led implementation).

We interpret these results as suggesting that reciprocity and pro-sociality is lower in channel construction groups relative to control, plausibly indicative of increased underlying conflict due to channel construction. Furthermore, referendum groups appear to have lower conflict within the command area as suggested by lower gap in their contributions between head and tail end farmers relative to data-driven groups. We do not find any strong evidence of local community-led implementation mitigating conflict within the command area.

5.4 Referendum as an (imperfect) Coordination Mechanism

Farmers liked voting, thought it was free, fair, and transparent. However, not everyone voted, particularly those that would have benefited most at baseline. Moreover, the resulting allocation of irrigation water is regressive. It is plausible that elites like large landholders could have influenced the consequences of the referendum, both in terms of who shows up to vote and also how farmers vote. It is also plausible that elites use referendum as a coordination device to minimize conflict and make side payments. We acknowledge that this interpretation is not conclusive and that additional research can shed light on the role of elites in conflict management in the context of managing common pool resources like irrigation infrastructure.

6 Discussion and Conclusion

In this paper, we show that there are advantages in learning about each community member's private preferences before the start of any development project. The standard approach has been that of centralized intervention to overcome coordination frictions. On the other hand, there are examples of many local institutions that exist to resolve this problem without centralized interventions. Many scholars and practitioners champion these latter approaches given the advantages of local communities in having information about their own preferences and being better in monitoring and enforcing contracts in principal-agent problems. However, such institutions are not common and we know very little about when and why successful cooperation arise.

We design an experiment with different aspects of an institutional design to construct field irrigation channels, where the status quo indicated Pareto inferior equilibrium of few functioning channels even when community members (farmers) desired more. We compare surface irrigation channel interventions with a budget-neutral control group, where no field channel construction was facilitated to identify the existence of collective action problem. Additionally, we vary whether local communities are involved in choosing the location of an irrigation channel through referendum voting and also whether these communities have agency in constructing the channel themselves. Importantly, we collected detailed production and private willingness to pay for each possible channel location at baseline. With this information, we were able to identify a data-driven benchmark that identified a location that maximized the aggregate value from surface irrigation across constituent farmers to compare against the outcome of a referendum.

The key implications of our study are the following: First, collective action and coordination problems lead to under-provision of commons such as irrigation infrastructure provision and maintenance. External facilitation of different forms can enable communities move out of this trap to a better equilibrium at a much lower cost to them. Second, referendum is not an ideal approach to identify social choice - the channel selected via voting diverges from the data-driven benchmark in close to half of the referendum villages with productivity implications. This could be because of poor turn-out and/or plausible strategic behavior among voters. Furthermore, subsequent allocation of irrigation water within referendum villages are regressive - farmers with lower marginal value get more irrigation than those with higher marginal value. In contrast, the data-driven benchmark of selecting aggregate value maximizing channel outperforms. Third, involving local communities in the construction has no substantial effect on either water allocation or other economic outcomes. To the contrary, this may delay the project timeline and increase visits by external facilitators to ensure timely completion.

Thus, for a policy-maker, it is important to learn about individual-level preferences since eliciting social choice is hard. The next step in this research agenda would be to collect data on private willingness to pay at scale. With modern technology and individual-level administrative data, this should be plausible. On the other hand, the actual implementation of development projects could be centralized as there are no immediate advantages that we could detect at our timescale. It is plausible that giving agency to local communities could strengthen their capacity in the long run, which remains an open question for future research.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?*," *The Quarterly Journal of Economics*, 10 2022, *138* (1), 1–35. 17

Balán, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel, "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo," *American Economic Review*, March 2022, *112* (3), 762–97. 2, 6

Bardhan, Pranab, "Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communities in South India," *Economic Development and Cultural Change*, 2000, *48* (4), 847–865. 5

__ **and Dilip Mookherjee**, "Decentralisation and Accountability in Infrastructure Delivery in Developing Countries*," *The Economic Journal*, 2006, *116* (508), 101–127. 2

Beath, Andrew, Fotini Christia, and Ruben Enikolopov, "Direct democracy and resource allocation: Experimental evidence from Afghanistan," *Journal of Development Economics*, January 2017, *124*, 199–213. 6

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, May 2014, *28* (2), 29–50. 18

Bó, Pedro Dal, Andrew Foster, and Louis Putterman, "Institutions and Behavior: Experimental Evidence on the Effects of Democracy," *American Economic Review*, December 2010, *100* (5), 2205–2229. 2, 7

Casey, Katherine, "Radical Decentralization: Does Community-Driven Development Work?," *Annual Review of Economics*, 2018, *10* (1), 139–163. 2, 6

__ **, Rachel Glennerster, and Edward Miguel**, "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*," *The Quarterly Journal of Economics*, November 2012, *127* (4), 1755–1812. 6

Chambers, Robert, *Rural development: putting the last first*, London ; New York: Longman, 1984. 2

Hayek, F. A., "The Use of Knowledge in Society," *The American Economic Review*, 1945, *35* (4), 519–530. Publisher: American Economic Association. 1

Kerwin, Jason, Nada Rostom, and Olivier Sterck, "Striking the Right Balance: Why Standard Balance Tests Over-Reject the Null, and How to Fix it," August 2024. 20

Khwaja, Asim Ijaz, "IS INCREASING COMMUNITY PARTICIPATION ALWAYS A GOOD THING?," *Journal of the European Economic Association*, 2004, *2* (2-3), 427–436. 6, 7

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119. 18

Mansuri, Ghazala and Vijayendra Rao, "Community-based (and driven) development : A critical review," Policy Research Working Paper Series 3209, The World Bank February 2004. 2

__ **and** __ , *Localizing Development: Does Participation Work?*, World Bank Publications, 2013. 2

Montero, Eduardo, "Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador," *Journal of Political Economy*, 2022, *130* (1), 48–93. 5

Mookherjee, Dilip, "Political Decentralization," *Annual Review of Economics*, 2015, *7* (1), 231–249. 1

Muralidharan, Karthik, "With rural economy as leading theme of polls, the policy options to address farm crisis," March 2019. 6

Oates, Wallace E., "An Essay on Fiscal Federalism," *Journal of Economic Literature*, 1999, *37* (3), 1120–1149. 2

Oates, W.E., *Fiscal Federalism* Harbrace series in business and economics, Harcourt Brace Jovanovich, 1972. 2

Olken, Benjamin A., "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia," *American Political Science Review*, 2010, *104* (2), 243–267. 6

Ostrom, Elinor, *Governing the Commons: The Evolution of Institutions for Collective Action* Political Economy of Institutions and Decisions, Cambridge University Press, 1990. 2

_ **and Roy Gardner**, "Coping with Asymmetries in the Commons: Self-Governing Irrigation Systems Can Work," *Journal of Economic Perspectives*, December 1993, *7* (4), 93–112. 2

Putterman, Louis, "On optimality in collective institutional choice," *Journal of Comparative Economics*, 1981, *5* (4), 392–403. 5

Qian, Nancy, "Making Progress on Foreign Aid," *Annual Review of Economics*, 2015, *7* (Volume 7, 2015), 277–308. 2, 6

Rao, Manaswini and Ashish Shenoy, "Got (clean) milk? Organization, incentives, and management in Indian dairy cooperatives," *Journal of Economic Behavior & Organization*, 2023, *212*, 708–722. 6

Rodell, Matthew, Isabella Velicogna, and James S. Famiglietti, "Satellite-based estimates of groundwater depletion in India," *Nature*, August 2009, *460* (7258), 999–1002. 6

Ryan, Nicholas and Anant Sudarshan, "Rationing the Commons," *Journal of Political Economy*, January 2022, *130* (1), 210–257. Publisher: The University of Chicago Press. 6

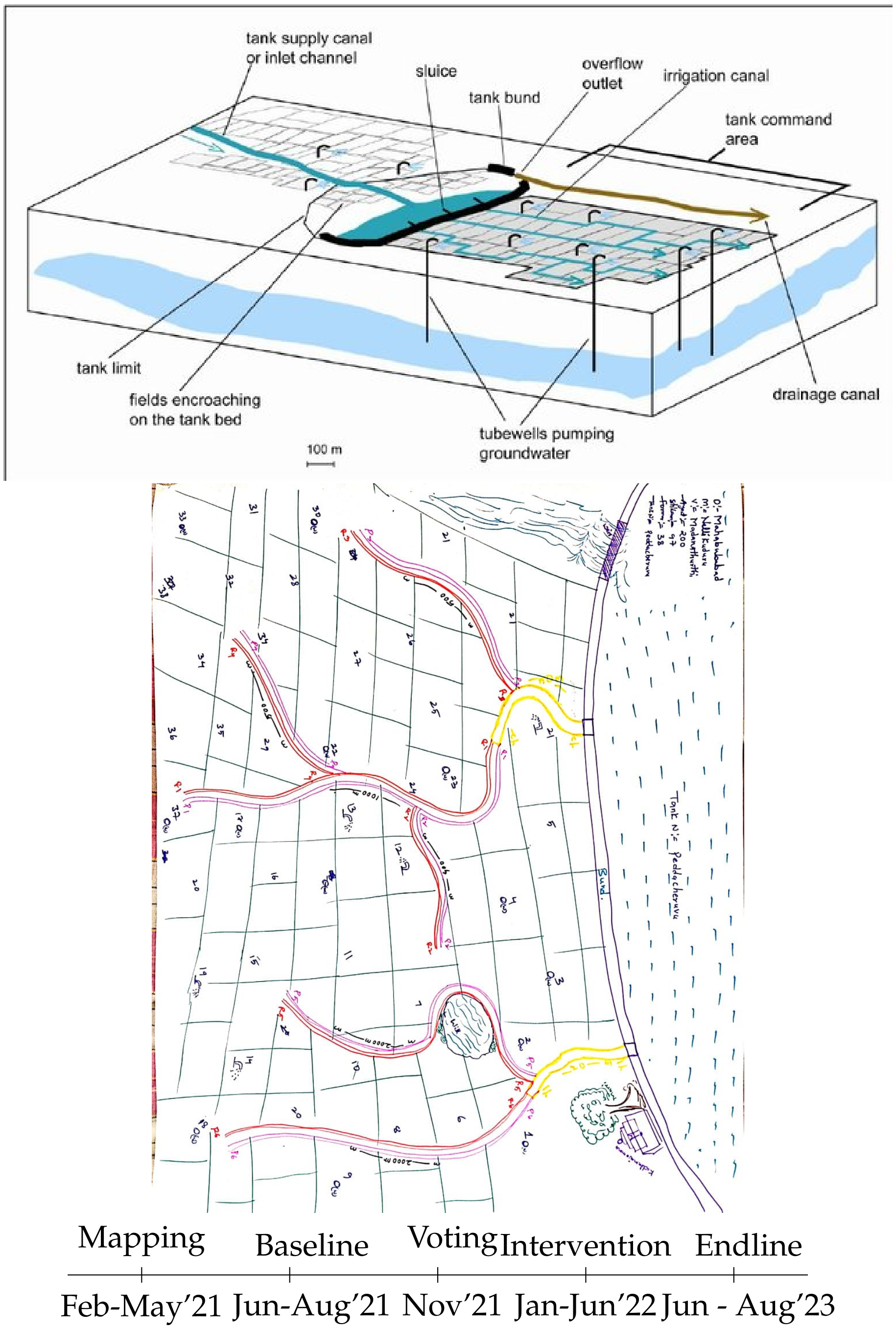
Sekhri, Sheetal, "Public Provision and Protection of Natural Resources: Groundwater Irrigation in Rural India," *American Economic Journal: Applied Economics*, 2011, *3* (4), 29–55. 6

Wade, Robert, "The management of irrigation systems: How to evoke trust and avoid prisoner's dilemma," *World Development*, April 1988, *16* (4), 489–500. 5, 8

Wong, Susan and Scott Guggenheim, *Community-Driven Development: Myths and Realities* 05 2018. 2

7 Figures

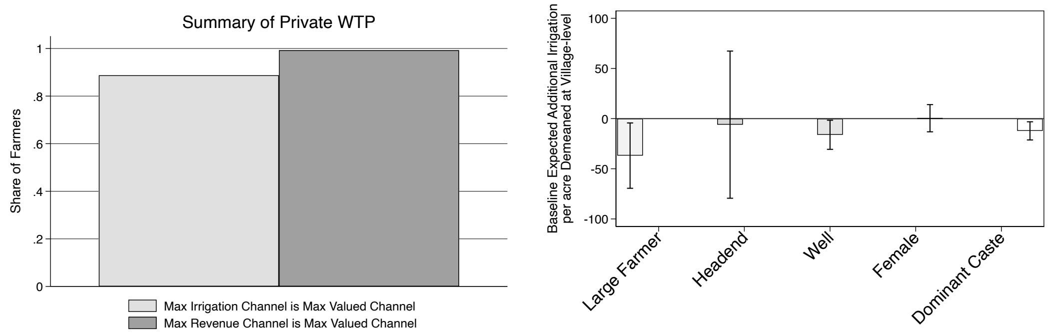
Figure 1: Schematic of Tank Irrigation



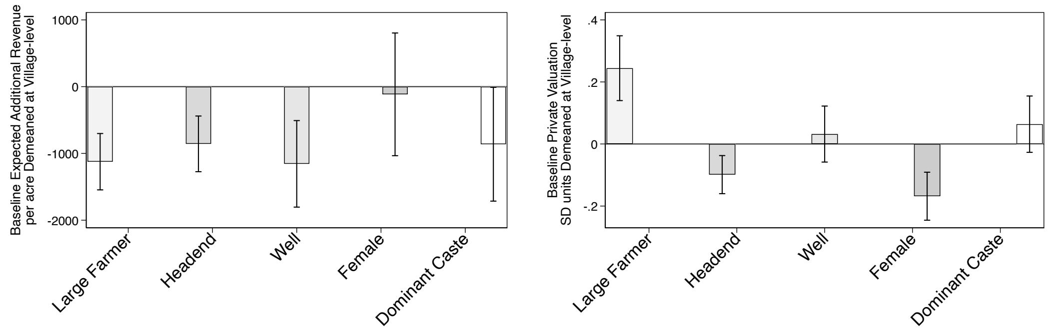
Notes: Image source top panel from Mathevet, et al, 2020.The schematic map in the bottom panel provides a map of feasible locations of channels, among which one was selected either via a data-driven selection mechanism, or via referendum, vs. none in control villages.

Figure 2: Baseline Location Choice and Private Willingness to Pay by Farm Characteristics

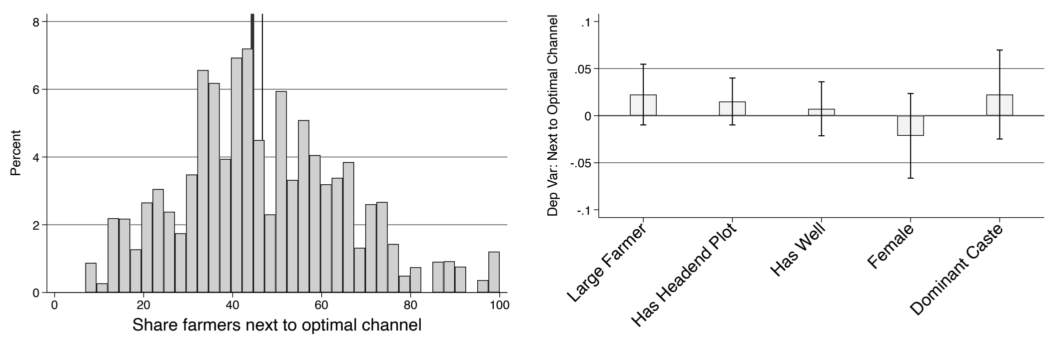
Panel A:



Panel B:



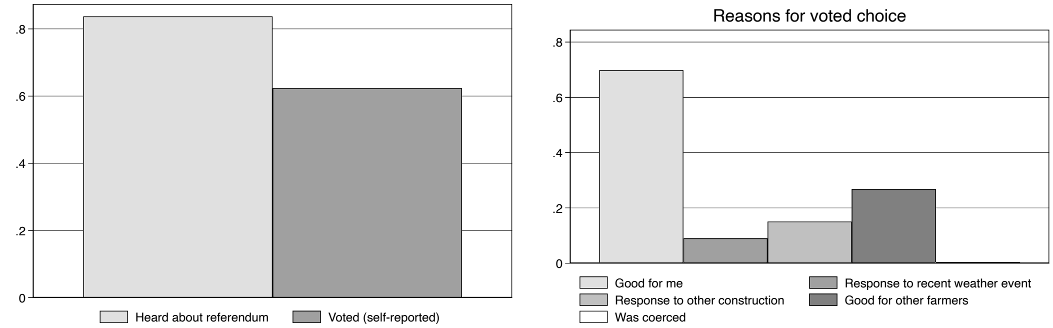
Panel C:



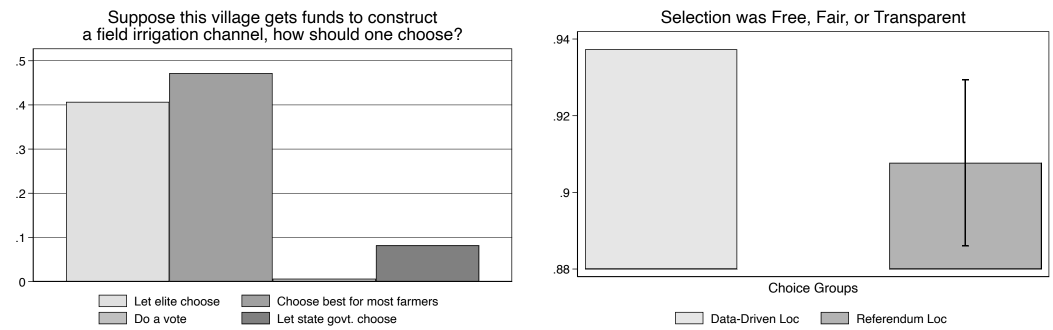
Notes: The figures above present the baseline stated willingness to pay (WTP) elicited for each feasible location for the channel and choice over how the channel location should be selected during private, in-person surveys across all study villages before the random assignment of the various interventions. In the elicitation, we collected additional days of irrigation anticipated from each feasible channel (that was later put on the ballot in a random fraction of villages), additional revenue earned from increased irrigation, and the final stated willingness to pay (private valuation). We first corroborate whether the location with maximum stated WTP also corresponds with the location resulting in maximum additional days of irrigation and maximum additional days of revenue (left figure, Panel A). We present the deviation from average values at the village-level by specific, pre-specified characteristics of farmers - whether they have a functioning well on their plot, whether their plot is in the head-end location, and whether they are large landholder (Panels A and B). Finally, Panel C presents the distribution of farmers along the aggregate surplus maximizing channel and whether their demographics correlate with this location. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure 3: Channel Selection

Panel A:



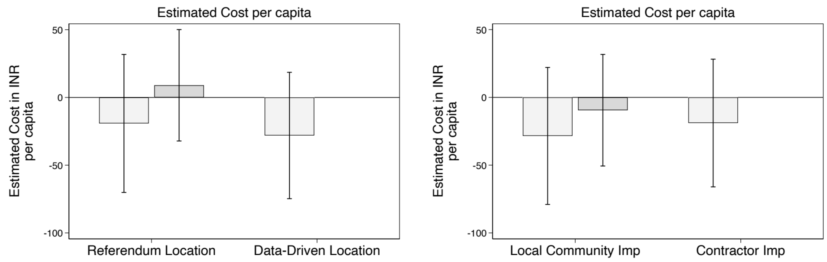
Panel B:



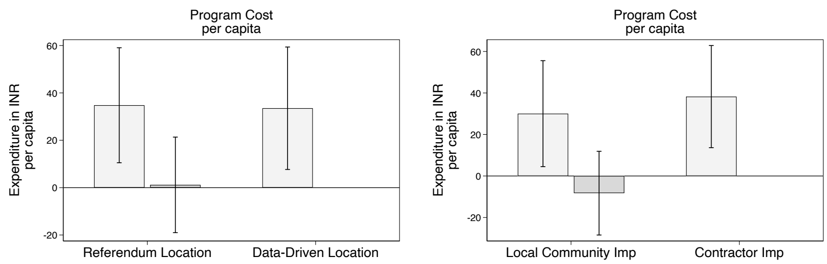
Notes: Sample includes only data-driven and referendum groups and excludes control group where no cash was distributed yet. Panel A presents descriptive statistics around implementing the referendum process in referendum groups (92 villages). Panel B presents response to how the location should be chosen at baseline (left) and sample farmers' opinion whether the selection was free, fair, or transparent after the channel location was selected and announced (right). Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure 4: Program Costs

Panel A:



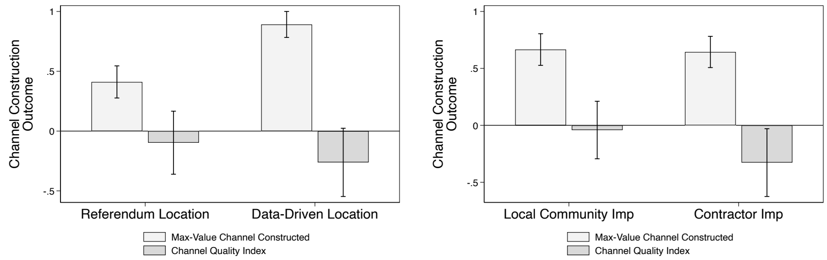
Panel B:



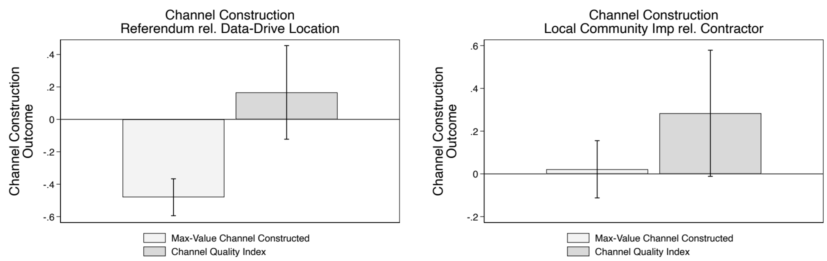
Notes: The figures examine program costs by different experimental groups. We normalize all cost in terms of per capita cost. Panel A presents per capita estimated cost at baseline and Panel B presents per capita realized cost. Lighter bars are differences relative to the control group whereas the darker bars are differences between choice and implementation groups, respectively. Error bars represent 95% confidence intervals. All specifications include randomization strata fixed effect. We report heterogeneity robust standard errors.

Figure 5: Channel Construction

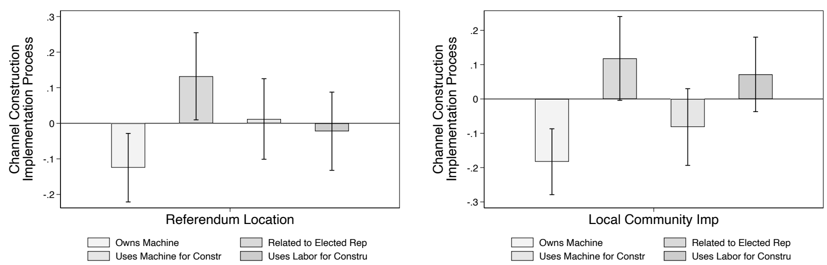
Panel A:



Panel B:

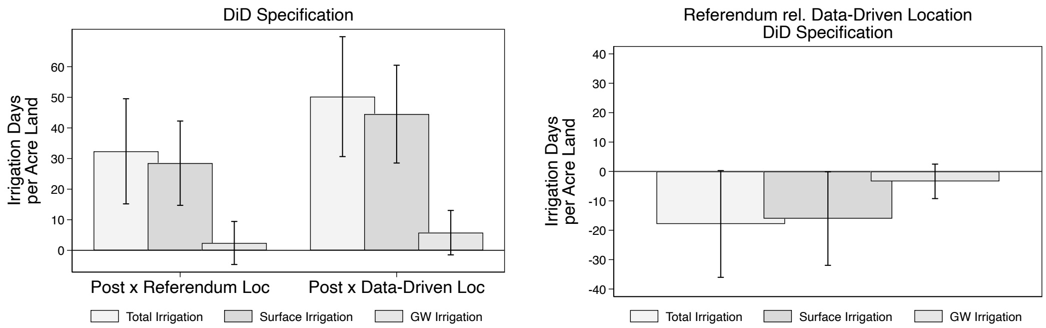


Panel C:



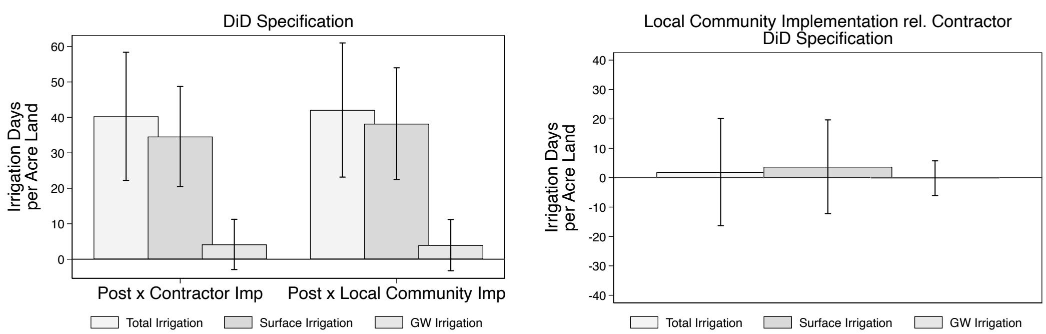
Notes: The figures examine channel construction outcomes during and at endline. We examine various process outcomes including whether channel with maximum aggregate valuation gets constructed, mode of implementation of the construction (who implements and how), and conditional on construction, what the quality of construction is. We generated an index for quality using various indicators such as the length of the channel, whether reaches tail-end, and whether local community made any additional repairs or modifications. We compare the construction outcomes both relative to control where 20% of the villages built some channel and relative to the respective choice or implementation groups, respectively. None of the control villages constructed the aggregate value maximizing channel as in the data-driven groups. In Panel C, we compare the mode of construction of the selected channels by choice (left) as well as implementation groups (rights). Control villages are excluded in these regressions as the details of the mode of construction is unobservable to us. Error bars represent 95% confidence intervals. All specifications include randomization strata fixed effect. We report heterogeneity robust standard errors.

Figure 6: Access to Irrigation: Referendum vs. Data-Driven Location Choice



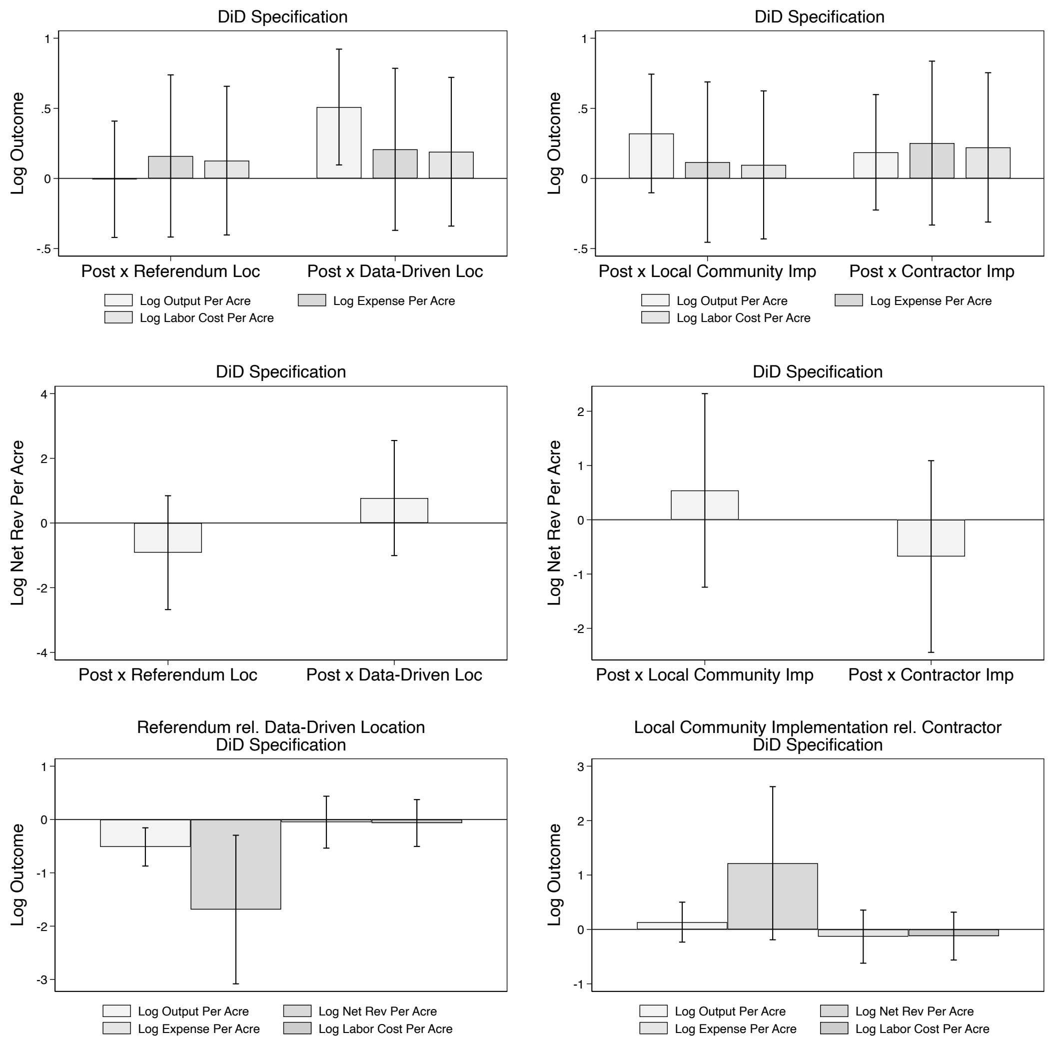
Notes: The figures above present the ITT estimates on number of days of irrigation adjusted for land area (left) and difference between referendum groups relative to data-driven location groups (right). The effects are substantial when compared to the budget-neutral arm where no channel is constructed in 80% of the control group. The increase in the average number of days of irrigation is consistent with the expected increase in about 30 additional days of irrigation anticipated at baseline. The differences in the estimates between referendum and data-driven location choice are also statistically significant at 10% and 5%, respectively. This difference represents a 15 day increase in irrigation per unit land under data-driven location choice representing an effect size of 0.2 standard deviation units. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure 7: Access to Irrigation: Contractor vs. Local Community Implementation



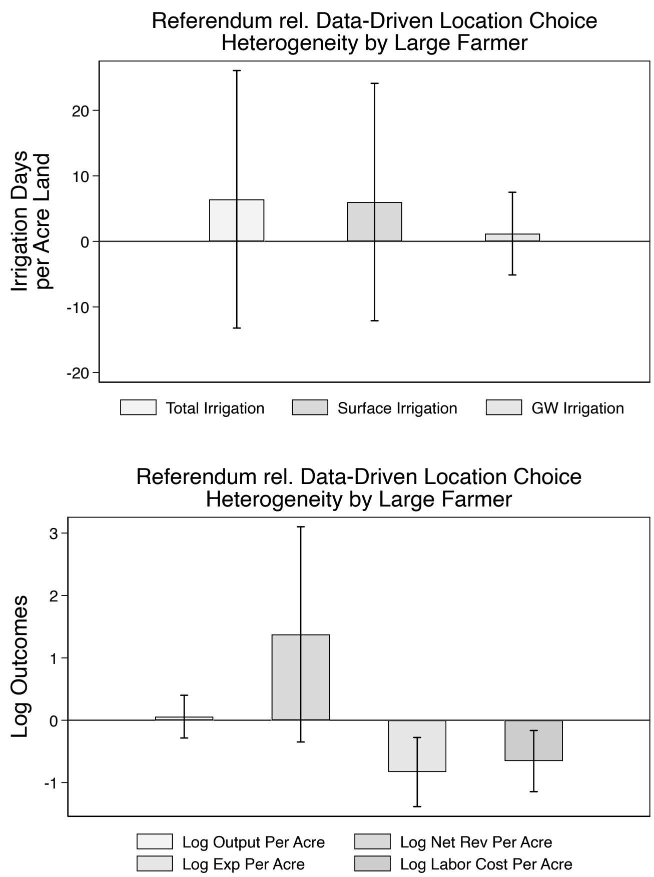
Notes: The figures above present the ITT estimates on number of days of irrigation adjusted for land area relative to control (left) and difference between local community-led implementation relative to contractor-led implementation (right). As seen before for the choice intervention, the effects of any implementation intervention are substantial compared to the budget-neutral control group. However, the differences between the specific type of implementation - contractor and local community-led implementation - are neither statistically significant nor meaningful in effect size. We are unable to reject any effect sizes lower than 0.02 standard deviation units (p-values>0.5). Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure 8: Production Outcomes



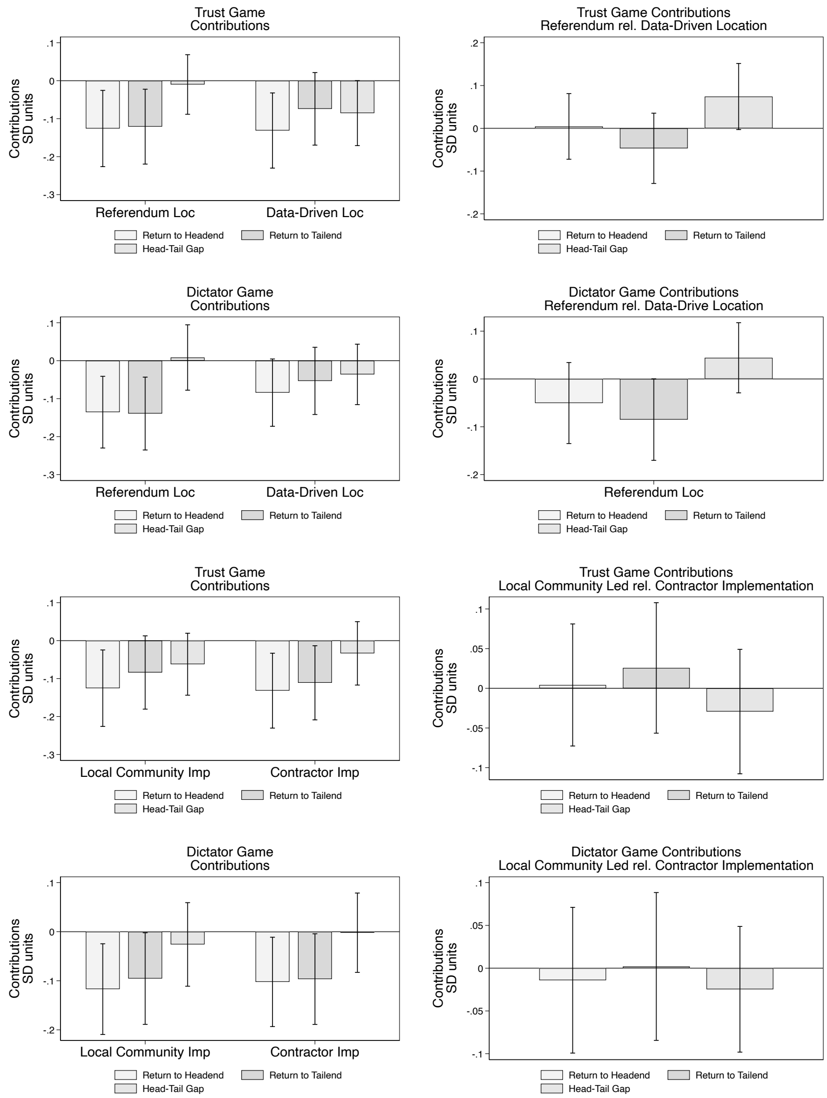
Notes: The figures above present the production outcomes per unit area - yield, net revenue, labor costs, total costs - across all the channel intervention groups relative to the budget-neutral control (top panel). The figure in the bottom panel compares referendum location choice villages with data-driven villages and local community-led implementation villages with contractor-led implementation villages, respectively. The differences in these estimates are statistically significant at 5% for output, representing a difference of 0.2 standard deviation units. The output and revenue effects represent overall income effects for the farmer households since a majority of the farmers cultivate only the parcel of land within the command area. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure 9: Referendum Generates Regressive Distribution



Notes: The figures above present the heterogeneous treatment effect estimates on number of days of irrigation and production outcomes by plot-size. Each coefficient is the coefficient on the triple interacted term, interacting the baseline characteristic by the 2x2 program choice experimental status. The comparison group is the set of small plots in data-driven location choice villages. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure 10: Conflict and Coordination



Notes: The graphs present respondent contribution to non-incentivized trust and dictator games respectively. The outcomes are game specific contributions at endline expressed in standard deviation units relative to the budget-neutral control group. Specifications controls for contributions in the games at baseline. The comparison/leave-out group is the budget-neutral control. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

8 Tables

Table 1: Summary Statistics and Balance Table

	(1) Referendum Loc Contractor Imp	(2) Referendum Loc Local Community Imp	(3) Data-Driven Loc Contractor Imp	(4) Data-Driven Loc Local Community Imp	(5) Control
			Village-Level		
Total Cultivated Area (Acres)	1426	1170**	1149	1576	1948***
	(1304)	(793)	(746)	(1213)	(3423)
Total Irrigated Area (Acres)	833	941	691	1007	842*
	(810)	(1738)	(636)	(1137)	(738)
Tank Area (Acres)	54	62	73	79	56*
	(60)	(83)	(82)	(74)	(47)
Tank Water Level (ft)	18***	15	13	12**	15
	(17)	(8)	(9)	(7)	(6)
No. Farmers	40**	44	44	43	45
	(10)	(14)	(21)	(13)	(14)
Total No. Channels	4	5	5	5	5
	(2)	(2)	(2)	(3)	(3)
Percent Not Used	47	47	47	44	41
	(30)	(30)	(28)	(26)	(25)
No. Feasible Locations	3	3	3*	3	3
	(1)	(1)	(1)	(1)	(1)
No. Villages	48	48	48	48	48
p-value (Joint Test)	0.003	0.64	0.11	0.11	0.002
			Farmer-Level		
Area (Acres) Cultivated by Farmer	3.55	3.74	3.61	3.53	4.07*
	(3.8)	(4.35)	(4.46)	(3.59)	(4.7)
Additional Days of Irrigation	27.04	29.21	31.45	35.44**	26.66
	(40)	(56.1)	(71.2)	(180)	(42.2)
Additional Revenue (INR)	4896.85**	5415.94	5923.20	5161.25*	5238.10
	(10616)	(15783)	(15957)	(8390)	(9258)
Stated WTP (INR)	4922.92**	4029.41*	4586.51	4809.45	4615.47
	(11889)	(5147)	(8534)	(8089)	(8356)
No. Farmers	1250	1244	1257	1278	1262
p-value (Joint Test)	0.05	0.19	0.77	0.07	0.29

Notes: The above table presents the average and standard deviation (in parenthesis) of each of the characteristics mentioned in the rows by the experimental groups in columns. Stars indicate whether specific values are determine assignment to specific groups. Although we used random number generator to assign each of the 240 villages to 5 experimental groups, it is plausible that some variables may not be "balanced" due to sample size restrictions and stratification. The last row indicates the p-value of joint test of equality of all the baseline characteristics determining specific assignment to an experimental group. All regressions include fixed effect of the randomizing strata (district). Standard errors are heteroskedastically robust for village-level regressions (Panel A) and clustered by village for farmer-level regressions (Panel B). Standard pair-wise t-tests for each variable across the 5 experimental groups are significant at 5% only for tank water level between data-driven local implementation relative to control and referendum contract implementation group relative to all data driven groups. None of the farmer-level baseline variables are significant in pair-wise comparison across the 5 experimental groups.

Table 2: Baseline-level of Coordination and Conflict

	(1) Referendum Loc Contractor Imp	(2) Referendum Loc Local Community Imp	(3) Data-Driven Loc Contractor Imp	(4) Data-Driven Loc Local Community Imp	(5) Control
Do you share (either take or give) water for irrigation	0.59	0.55	0.59	0.56	0.63*
	(0.49)	(0.5)	(0.49)	(0.5)	(0.48)
Coordination at Village	0.49	0.47	0.48	0.48	0.49*
	(0.31)	(0.31)	(0.3)	(0.31)	(0.3)
Coordination at Command Area	0.51	0.49	0.50	0.51	0.51
	(0.31)	(0.31)	(0.3)	(0.3)	(0.29)
Have you had tensions with your plot neighbors	0.07	0.07	0.07	0.09	0.09
	(0.25)	(0.26)	(0.25)	(0.29)	(0.28)
Is there any formal or informal water sharing rules	0.14	0.14	0.17	0.21*	0.15
	(0.35)	(0.35)	(0.38)	(0.41)	(0.36)
No. Villages	48	48	48	48	48
No. Sample Farmers	1246	1230	1229	1257	1247
p-value (Joint Test)	0.72	0.52	0.93	0.13	0.17

Notes: The above table presents the average and standard deviation (in parenthesis) of each of the characteristics mentioned in the rows by the experimental groups in columns. Stars indicate whether specific values are determine assignment to specific groups. Although random number generator was used to assign each of the 240 villages to 5 experimental groups, it is plausible that some variables may not be "balanced" due to sample size restrictions. The last row indicates the p-value of joint test of equality of all the baseline characteristics determining specific assignment to an experimental group. All regressions include fixed effect of the randomizing strata (district). Standard errors are clustered by village for farmer-level regressions.

Table 3: Characteristics of Selected Channel in Referendum Villages

	(1) Next to Chl w/ Highest Votes	(2) Next to Max-Valued Chl	(3) Next to 2nd Max-Valued Chl	(4) Voted Referendum	(5) Voted Max Val Chl	(6) Voted Private Max Chl
Large Plot	-0.0177	-0.0181	-0.00853	0.0474**	0.0312	0.0346
	(0.0283)	(0.0308)	(0.0299)	(0.0232)	(0.0205)	(0.0218)
Headend	-0.0305	0.0293	-0.0255	0.0611**	0.0300	0.0119
	(0.0296)	(0.0297)	(0.0290)	(0.0252)	(0.0205)	(0.0202)
Has Well	0.00815	0.00606	0.0142	0.0207	-0.00672	-0.0265
	(0.0290)	(0.0294)	(0.0268)	(0.0241)	(0.0242)	(0.0220)
Female Cultivator	0.0447	-0.0567	0.0269	-0.0537	0.0407	-0.0465
	(0.0455)	(0.0474)	(0.0505)	(0.0428)	(0.0325)	(0.0311)
Dominant Jati	0.0185	0.0224	-0.00575	-0.0261	-0.0197	-0.0270
	(0.0554)	(0.0500)	(0.0508)	(0.0516)	(0.0303)	(0.0306)
Is Influential	-0.000803	-0.000631	-0.0544	0.0245	0.0334	0.0323
	(0.0383)	(0.0354)	(0.0361)	(0.0278)	(0.0255)	(0.0276)
Next to Chl w/ Highest Votes				-0.00143	0.0218	0.109***
				(0.0250)	(0.0180)	(0.0217)
Next to Max Valued Channel				-0.0396	0.0478**	0.0139
				(0.0319)	(0.0233)	(0.0234)
Next to 2nd Max Valued Channel				-0.0706**	-0.0263	-0.0000215
				(0.0306)	(0.0216)	(0.0234)
Obs	2185	2185	2185	2185	2185	2185
Command Areas	96	96	96	96	96	96
FE	Village	Village	Village	Village	Village	Village
Mean	.36	.45	.33	.43	.17	.19
SD	.48	.5	.47	.5	.38	.39
Adj R2	0.139	0.110	0.0791	0.162	0.132	0.0797
Model F-stat	0.538	0.516	0.589	2.985	2.763	5.100

Standard errors in parentheses
* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: Rows are baseline demographic characteristics of the random sample of farmers or plots in the command area in 96 referendum villages. The Columns variables are as follows - dependent var in Col 1 is whether the sample farmer is along the channel with the highest votes, Col 2 is whether the sample farmer is along the max-valued channel, Col 3 is whether the sample farmer is along the second-max valued channel, Col 4 is whether sample farmer participated in the referendum, Col 5 is whether the sample farmer voted the max channel for the village, and Col 6 is whether the sample farmer voted for channel with highest value for themselves in private. Dep var is coded 0 in Col 4-5 if the sample did not participate in the referendum. Standard errors are clustered by village command area to account for any spatial correlation.

Table 4: Village-level Voting Patterns

	(1) Referendum Turnout	(2) Referendum Turnout	(3) Referendum Turnout	(4) Referendum Selects Max-Value Chl	(5) Referendum Selects Max-Value Chl	(6) Referendum Selects Max-Value Chl
Share Next to Max-Valued Chl	0.125	-0.00186		0.00300	-0.00217	
	(0.115)	(0.134)		(0.00295)	(0.00252)	
Share Next to 2nd Max-Valued Chl	0.271*	0.174		0.000850	-0.00314	
	(0.151)	(0.152)		(0.00343)	(0.00352)	
Share Next to Max Votes Chl		0.298**			0.0121***	
		(0.135)			(0.00156)	
Share Large Adj			-0.210			-0.00119
			(0.148)			(0.00373)
Share Headend Adj			0.0554			0.00224
			(0.104)			(0.00252)
Share Has Well Adj			0.0177			0.000287
			(0.0738)			(0.00172)
Share Female Adj			-0.0975			0.0120**
			(0.201)			(0.00529)
Share Dom Jati Adj			0.265**			-0.000733
			(0.102)			(0.00306)
Share Influential Adj			-0.480**			-0.00318
			(0.213)			(0.00499)
Obs	96	96	96	96	96	96
Mean	43.84	43.84	.44	.44	.44	.44
SD	21.98	21.98	.5	.5	.5	.5
R2	0.0393	0.105	0.126	0.0118	0.222	0.0561

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: The sample only includes referendum villages to examine village-level voting patterns. Col 1-3 report village-level turnout as the dependent variable. The dependent variable in Col 4-6 is whether the referendum selects the same channel as using the data-driven approach, i.e., using the aggregate value maximizing channel. The explanatory variables are share of farmers in the command area with the specific characteristics. For example, "Share Next to Max-Valued Chl" is the share of farmers who are along the aggregate value maximizing channel. Heteroskedasticity robust standard errors in parantesis.

Table 5: Administrative Data on Implementing the Intervention

	(1) No Data	(2) Overall Difficulty Implementation	(3) No. Visits Team	(4) Delay Payment Acknowledgement	(5) Delay Appointment	(6) Delay Construction	(7) Demand Additional Construction
Referendum Loc x Contractor Imp	-0.0432	-0.852**	-0.797	-0.0714	-0.0771	-0.0144	0.0423
	(0.0421)	(0.411)	(0.599)	(0.0778)	(0.0848)	(0.0842)	(0.0555)
Referendum Loc x Local Community Imp	-0.0344	0.0684	0.120	-0.0325	0.118	0.0312	0.0461
	(0.0418)	(0.499)	(0.631)	(0.0816)	(0.0900)	(0.0836)	(0.0530)
Data-Driven Loc x Contractor Imp	-0.0365	0.245					
	(0.0395)	(0.454)					
Data-Driven Loc x Local Community Imp	-0.0343	-0.0211	0.666	0.0829	0.0128	0.143*	0.106*
	(0.0414)	(0.453)	(0.704)	(0.0860)	(0.0924)	(0.0858)	(0.0570)
Obs	240	225	186	186	186	186	186
FE	District	District	District	District	District	District	District
Control Mean	.09	3.85	5.42	.33	.56	.51	.16
Control SD	.29	2.8	3.57	.48	.5	.51	.37
Ballot=No Ballot	.9	.18	.14	.1	.8248	.3094	.8356
Top Down=Bottom Up	.84	.38	.08	.28	.107	.1297	.1898

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: The above table presents the results on governance outcomes in the process of providing the intervention (channel construction) itself. Column 1 records any differential attrition in the administrative data, which we find none. Column 2 is a self-reported measure of difficulty in completing the intervention (channel construction in all treatment and cash disbursal in control) as reported by the field research team coordinating the interventions. Column 3 records the number of visits by the field research teams to enable completion of the construction. Columns 4-6 record time delays in acknowledgement/response from the communities in confirming receipt of payment, appointment for our visit, and the final completion of the construction. Column 7 records whether the community asked for additional channels to be constructed from the research team. Columns 3-7 do not include the budget-neutral control since the administrative data was about channel construction facilitation provided as part of our research design. Therefore, for these columns, the leave-out group is the No-Choice Top-Down group. Standard errors are clustered by village and the specifications include randomization strata fixed effect. Disclaimer: We could not comply with any of the study village community request to build additional channels due to maintaining sanctity to research design and to comply with our project budget. We politely convinced the communities of our study objectives during post-intervention debrief.

Table 6: Status of Channel Constructed at Endline

	(1) No Audit Endline	(2) Channel Status Endline	(3) Construced Length	(4) Reach Tailend
Referendum Loc x Contractor Imp	-0.000165	0.601***	294.2***	0.642***
	(0.00532)	(0.0801)	(28.23)	(0.0809)
Referendum Loc x Local Community Imp	0.0187	0.646***	292.3***	0.646***
	(0.0192)	(0.0743)	(29.62)	(0.0721)
Data-Driven Loc x Contractor Imp	0.0190	0.517***	300.5***	0.441***
	(0.0195)	(0.0837)	(31.45)	(0.0838)
Data-Driven Loc x Local Community Imp	-1.63e-18	0.625***	299.5***	0.667***
	(0.00526)	(0.0779)	(30.99)	(0.0710)
Obs	240	238	238	238
FE	District	District	District	District
Control Mean	0	.19	60.42	.15
Control SD	0	.39	132.07	.36
Ballot=No Ballot	.99	.3	.7607	.1
Top Down=Bottom Up	1	.13	.9488	.0369

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: The table presents village-level data on random audits by the research team. Column 1 shows any differential attrition in the site visits, which we find none. Columns 2-4 record specifics of the constructed channel, including whether they are functional (Col 2), length (Col 3), and whether they reach the tail-end within the command area (Col 4). Columns 5-7 are about connections to the main village reservoir/tank. We incorporate heterogeneity robust standard errors and include randomization strata fixed effect.

Online Appendix

A Experiment Protocol and Scripts

A.1 Pre-baseline Activities

Village Sampling Frame: We used the government tank census data to identify villages with tanks above 100 acres command area. We called the respective Sarpanch/VRO to check whether there is tank based surface irrigation in any of the seasons (Kharif or Rabi). We recorded all this information in an excel file (data to be entered in exactly in the format provided). Details to select a village into the sampling frame will be based on the presence of surface irrigation after on-site verification by research team members.

Preparation of Command Area Maps A team of 2 research assistants sought appointment from the village sarpanch and VRA for the mapping exercise. The team conducted a focus group in the village with sarpanch, VRA, and a group of farmers representing head, middle, and tail regions, to draw a schematic map of the tank, existing irrigation channels, and identified potential locations for constructing field channels. Using the map of tank command area, they marked the proposed locations of the potential new channels (for example, if it is a sub-channel, they marked it so on the map) to be constructed. For reference, they also included landmarks on the maps for easy identification of the location. Next, they estimated the cost of constructing each of the potential channel. For example, if one channel would need 4 male and 4 female laborers for 2 days of work, then the cost of the construction would be the total labor cost as per the local prevailing wage rate.

Script: *We are interested in improving farmer welfare and particularly interested in understanding agricultural practices in this village. We want to understand how command area farmer community decide on using water from the tank (mention tank name) for surface irrigation. We will ask you and this group a few questions about tank use and make a schematic map of the tank and its command area.* **Draw a schematic map.**

Now, looking at this map, can you identify 3-4 locations where a new channel or sub-channel can be constructed to improve surface irrigation from the tank. Mark the location on the map. We will also be approaching the farmers to understand their cultivation practices and identify their preference for the location of the potential channel to be constructed. **Collect farmer listing data.**

Once we have completed data collection, we will either identify which channel to construct ourselves or we will hold a secret ballot so that the community can choose which channel should be constructed. Next, we will either *construct the channel ourselves or provide funds to your community or farmers directly to construct yourselves. We will collect data on irrigation use and agriculture production once every growing season for the next 1-2 years. We thank you for your cooperation and we will continue to be in touch through this study.*

Farmer Sampling Frame: We enumerated of all farmers within the tank command area, by section, i.e. Head, Middle, and Tail. For each farmer, we recorded their location (head/middle/tail) and whether they have a functioning bore-well on their plot currently. We also recorded their phone numbers for contacting for the baseline survey.

A.2 Baseline Private WTP elicitation

Greetings! We are from J-PAL South Asia organization, representing researchers who are interested in understanding and addressing the problems concerning effective usage of tank water for irrigation to increase agricultural production, yield, and ultimately farmer welfare. Towards this goal, we have come to you with X number of proposals to build channels/mud trenches (but not permanent concrete ones) in this tank's (mention name) command area. The locations of these channels were proposed by sarpanch, VRA, neeretigar, and asamis from this command area during a village meeting we organized in your village Y days ago. [ENUMERATOR Instruction: Show the map with the location and the extent of the channels.]

We want to first understand whether or not any of these channels are valued by the farmers more than the current status quo and if so, which among these have the highest value. Specifically, we want to identify one channel that generates the highest value across all farmers in the command area and whether this total value is more than the cost of constructing or repairing the particular channel. Please note that by valuation, we don't mean that you have to pay this amount. The valuation is how much would you gain (in money terms) from cultivating your plot in the command area if the channel was constructed or repaired.

Consider the following example. Under status quo of channels in disrepair, you face uncertain access to surface irrigation. Because paddy requires a lot of water, this can affect the yield. Let's assume that the revenue you earn from selling paddy under status quo is INR 30,000. Now suppose that the channel is built or repaired such that you now have more reliable access to surface irrigation. This increases the paddy yield and the associated revenue to INR 35,000. This means that the channel's value is upto INR 5000 (35000 – 30000). [ENUMERATOR Instruction: Check if this process of valuation is clear to the farmer].

If the total value of the highest-valued channel across all farmers is less than the cost of building or repairing that channel, then we may not construct or repair any channel in your village at all. [ENUMERATOR Instruction: Mention the approximate cost for each of the channel, mentioning the length and showing it on the map.]

Remember, you will not be required to pay this amount, but you need to truthfully mention your honest and independent valuation. We will add your valuation to the valuation provided by other farmers to identify which channel generates the highest value across all farmers in the command area. If you are not being truthful and quote a very low value much less than your true valuation, others may also do so and in such a situation no channel will be constructed or repaired.

Our organization will be facilitating the construction or repair of one of these channels through a cash subsidy of up to INR 10,000. We will either provide this money to this village (someone recommended by the initial group we met including sarpanch) to construct yourselves or we may construct ourselves.

Regarding which channel will be selected, we will either hold a secret ballot voting in the village later or we will decide the channel ourselves. If we decide ourselves, there will be no voting.

Channel construction or repair is going to cost money and/or labor. Therefore, you will only be asked to contribute the difference between how important the channel is for you relative to other farmers. Similarly, other farmers will also be asked to contribute the difference between how important the channel is for them compared to all other farmers in the command area. What this means is that you will only be asked to contribute if the highest-valued channel will change if your valuation for the channel is dropped when calculating the total. It may be the case that dropping your valuation will not change which channel generates the highest value across all other farmers. In such a situation, you pay nothing. However, if dropping your valuation changes which channel generates highest value across all other farmers, then, you will be asked to pay the difference between the value of the new highest-valued channel and the total valuation from all other farmers for the current highest-valued channel. We will further net out the subsidy we provide (INR 10000) from the total cost of construction and you will only need to contribute towards the additional cost of construction. You may contribute this share either in cash or volunteer to manage the construction/repair, as you deem fit. What this means is that under such a situation where the channel is more important to you relative to other farmers, you will need to take additional responsibility towards channel construction, either in terms of paying part of the cost or in kind through voluntary labor.

Note that if you artificially inflate your valuation higher than your true valuation, you may actually be required to pay some contribution which you need not if you state your valuation truthfully.

To understand this exercise better, consider the following example. Suppose there are three farmers in the command area – Ramu, Shamu, Bheemu. The village has proposed two channels that need to be constructed or repaired. Channel G1 costs INR 10500 and G2 costs INR 3000. Ramu, Shamu, Bheemu provide their valuations as shown in the table below:

Farmer	Channel: G1	Channel: G2	Payment (without subsidy)	Payment (with subsidy)
Ramu	5000	6000	0	0
Shamu	8000	0	1000	500
Bheemu	3000	3000	0	0

Clearly channel G1 has the highest valuation (16000 relative to 9000 for G2). Suppose this channel is selected for construction/repairs. If we remove Ramu's valuation, G1 is still the highest valued channel (valued at 11000 compared to G2 valued at 3000), so he pays nothing. If we remove Bheemu's valuation also, G1 continues to be the highest channel (valued at 13000 compared to G2 valued at 6000), so he also pays nothing. However, if we remove Shamu's valuation, then the highest value channel is G2 (valued at 9000 compared to G1 without Shamu, which is now valued at 8000). So, the difference is 9000-8000 = 1000 and this is how much Shamu will have to pay if there is no subsidy. But, remember that we provide INR 10,000 whereas the cost of constructing G1 is INR 10,500. In this case, Shamu only has to pay 500 (1000 – 500). Now, Shamu may pay this in cash or may volunteer to monitor the construction process, based on what he is comfortable with.

[ENUMERATOR Instruction: Check if it is clear to the respondent that not all have to pay some cost for channel and those that may have to pay some cost do not pay anything close to their actual valuation of the channel.]

Instruction: Once the enumerator has ascertained the understanding of the respondent, record their valuations for the different proposed channels.

A.3 Cost-Neutral Control Protocol

Field managers distribute per-farmer average cost of channel construction in the intervention villages. They used the following script at the time of cash disbursal. *Greetings! We are from J-PAL South Asia at IFMR, representing researchers who are interested in understanding and addressing the problems concerning effective usage of tank water for irrigation to increase agricultural production, yield, and ultimately farmer welfare. We collected some data from you a few weeks ago.*

We are now providing you with $\approx$ INR 300 (we gave the computed value, which is approximated 300 per farmer) as well as to other study farmers in this command area. You may decide to use this money together with other farmers to build an irrigation channel or decide to use this for any other personal use. There are no conditions attached to this cash.

A.4 Ballot Voting Protocol

Team size in each district: 2 surveyors + 1 supervisor No. of voting days per village: 4 days Items needed: Ballot box, ballot cards, maps, farmer list to guide and track distribution Targeted group: Universe of all command area farmers as per our initial listing data (in villages with bottom-up choice)

1. Day 0

   (a) Supervisor to call the Sarpanch to inform that we would start a ballot voting exercise in the village the next day for the choice of channel to be constructed.

   (b) Before starting the discussion about voting exercise, enquire about the covid situation in the village to ensure that it doesn't breach the protocol of <2% active cases in the village (this protocol will be updated in line with learnings from the early experience and how the covid situation in the state progresses).

   (c) Ask for permission to place the ballot box within the premises of the village's panchayat office for the next 4 days. If the gram panchayat office is located in another village or at a location that is not easily accessible, ask the Sarpanch to suggest a public property within whose premises the ballot box can be placed safely (for instance, any local govt office/school/ ration shop). Also ask the sarpanch to suggest a person for the box's safety who could take it in at 7 pm and place it out at 8 am for the next 4 days and ensure that the box is not tampered with.

   (d) Make sure that the ballot cards, farmer list and maps are ready.

2. Day 1

   (a) Reach the village by 9 am and place the ballot box at the decided location. Talk to the person suggested by the Sarpanch and convey the purpose and details of the project in a way that they understand the vitality of the voting exercise.

   (b) Place an A3 copy of the command area's map on the wall right next to where the ballot box is placed. Make sure that both ballot box and map are placed at a location where they are unaffected by rain/harsh weather. Post a photo of the ballot box and map taken using the timestamp app on the whatsapp group.

(c) After the ballot box is placed, go door-to-door to visit all the farmers on the list to invite them to vote. (This list includes all the farmers in the command area as per the initial listing exercise and not just the baseline sample). In SCTO, select all UIDs to whom the card was distributed.

(d) Convey all the necessary information as per the script below and hand over the ballot card and a map to all the farmers:

Greetings! We are from J-PAL South Asia at IFMR, representing researchers who are interested in understanding and addressing the problems concerning effective usage of tank water for irrigation to increase agricultural production, yield, and ultimately farmer welfare. We collected some data from you a few weeks ago. We now want to identify the location of the channel to be constructed through secret ballot – one farmer, one vote. I will now explain the different location choices of the channels and hand you a copy of the map. Please tick the channel that you prefer the most and drop the ballot card in the ballot box placed at location (mention the details) on or before the last date. Your vote will be anonymous. You may ask for assistance from us at the project contact number. After the vote, we will select the channel for construction as the one that receives highest votes. We will then proceed with the construction shortly thereafter.

3. Day 2: Open for voting between 8 am and 7 pm.

4. Day 3: Open for voting between 8 am and 7 pm. On this day, we would send the following reminder SMS to all farmers. *Greetings. Our team visited you last week and gave you a ballot card and a map. Your vote will help us decide which channel gets selected for construction/repair. Please make sure to drop your ballot card in the box by tomorrow. J-PAL South Asia*

5. Day 4: Visit the village and collect all the ballot cards from the box at 5:30 pm. After the cards are collected, leave the village immediately and count and upload the voting data on SurveyCTO (For each village, mark the chosen channel for each UID - NA if card is missing) after reaching home on the same day. Any discussion with the Sarpanch/the person assisting with overlooking the ballot box/anyone else in the village should only take place before the ballot cards are collected. Do not show the ballot cards to anybody. Let the Sarpanch know that the ballot box is being left behind in the village and can be used by them

A.5 Bottom-up Implementation Protocol

In addition to following the above-mentioned protocol following the choice intervention, we provide the following information in villages that were assigned to bottom-up implementation arm:

We will be providing the collected funds along with our contribution to Mr. Z from your village (mention the name), who a majority of you have unanimously appointed as the "Project Director". We will be providing 50% of the cost now to start the project by this date (mention the starting date), or when the works have started and the remaining 50% upon satisfactory completion of the sub-channel construction. We expect the sub-channel to be ready by this date (mention the end date).

B Data Description

Intermediate Outcome: Tank-Level Community Effort These include the first-stage measures of type of the public good (channel), including: (a) whether a channel is constructed, (b) status, including deviation, of the constructed channel relative to the initial proposed design (length, width, depth, geo-trace polyline) at multiple time periods post construction, (c) duration of construction, (d) local labor hours used for construction, (e) whether machinery is used to dig the channel, and (f) total realized cost of construction (including deviations from the budgeted cost). We will use a summary index measure combining all the above components of the first stage, in addition to examining these separately.

Intermediate Outcome: Farmer-Level Coordination These include: (a) complementary investments by command area farmers (effort cost incurred during construction and maintenance including voluntary labor contribution, monitoring costs, cash and material support) as a measure of local community effort, (b) farmer satisfaction with participating in decision-making and construction of the channel, (c) measures of trust between farmers with plots in different locations within the command area, (d) farmer-level voting outcomes (whether aligned with private elicitation, and whether voting influenced by others in the village), and (e) cost of participation in voting (time cost involved in voting) and implementation (hours worked beyond compensation).

Primary Outcome - Plot-Level Access to Irrigation These include: (a) surface water availability at sample plots in various locations (head, middle, tail) within the command area (extensive margin - whether receives, as well as intensive margin - number of days), and (b) extent of water-related conflicts (number of conflicts, number of farmers involved in the conflict).

Primary Outcome - Plot-Level Production We will examine total production (quantity produced) and total yields (qty per acre), in aggregate as well as on average across a random sample of command area farmers. Usually, farmers in the command area grow paddy since paddy requires irrigation. We will verify the crop grown so that the production outcomes are specific to crop grown.

An important outcome relating to our research design is the cost per unit increase in aggregate agricultural production output - cost per unit increase in quantity of output, cost per unit increase in yield.

Secondary Outcomes - Plot-Level Climate Change Mitigation and Adaptation Two important consequences of helping overcome coordination and collective action failures in the context of irrigation management are climate change mitigation and adaptation. First, reliance on surface water for irrigation minimizes extraction of groundwater resources, addressing climate change mitigation. We measure the extent of groundwater use at the sample plot-level, both on the extensive margin (number of plots with functioning well) as well as on the intensive margin (number of days). Second, increase in access to (whether sample plot has access to tank water through field irrigation channels) and the extent of (number of days) surface irrigation through rainwater harvested in the tanks addresses climate change adaptation to increased drought conditions (specifically, aiding cultivation during the dry season or wet seasons with low rainfall).

Secondary Outcomes - Plot-Level Production Cost and Revenue These include plot-level measures of cost and revenue from agricultural production: (a) sales revenue by crop-season, and (b) expenditure by crop-season and expense-type (i.e. labor, capital, fertilizers and input). Since we don't expect prices to change due to our treatment, these are secondary outcome measures.

Secondary Outcomes - Village-Level Agricultural Wages by Gender Since paddy is typically the main irrigated crop cultivated within the command area, the gender dynamics are particularly important to study given the substantial role played by female agricultural labor at various stages of paddy lifecycle. We collect detailed agriculture wage data, by gender and agricultural task.

C Figures

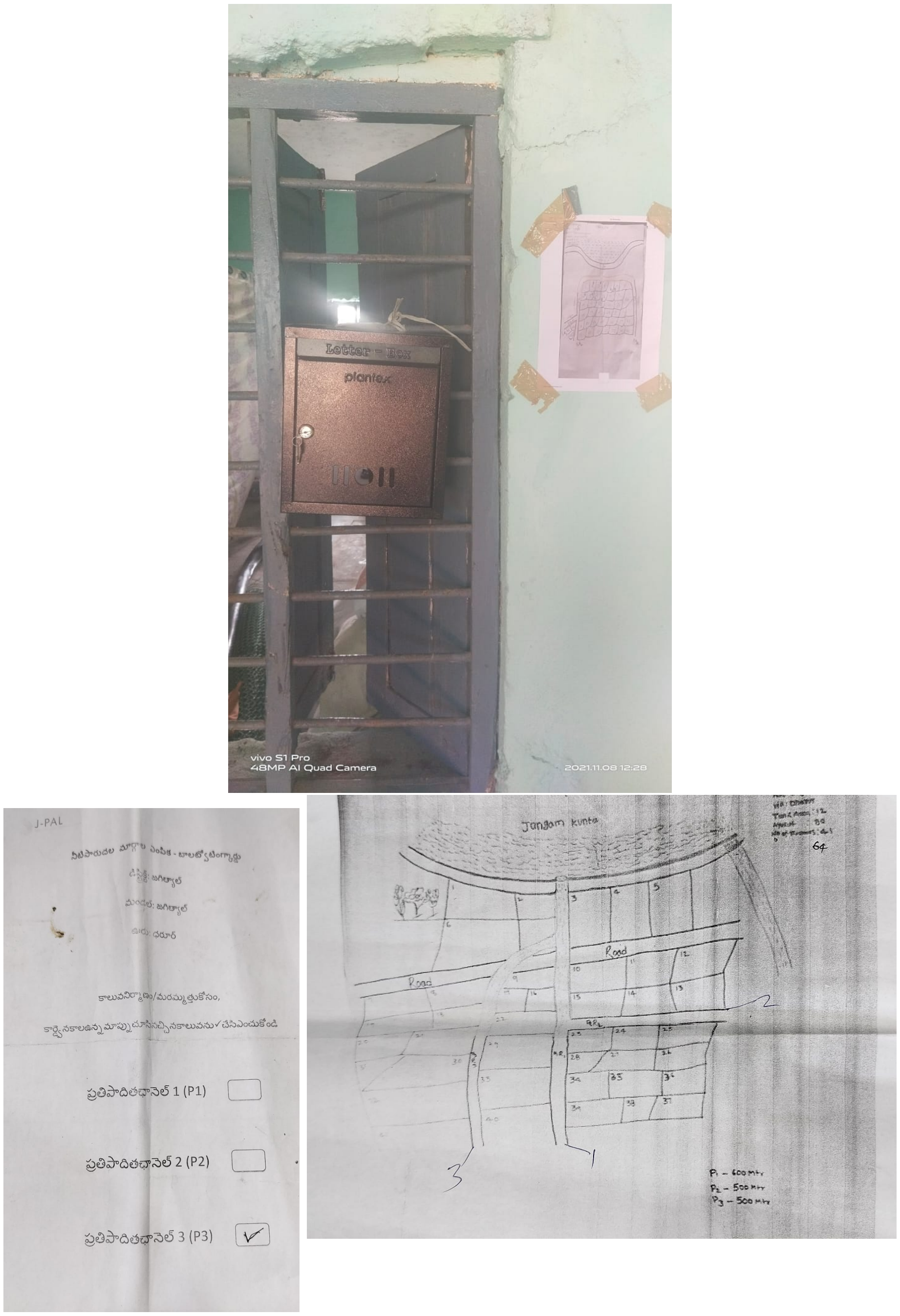
Figure A1: Images Before, During, and After Channel Construction

Notes: The images show the status of the command area before the intervention.

Figure A2: Long-Run Channel Status of Constructed Channel

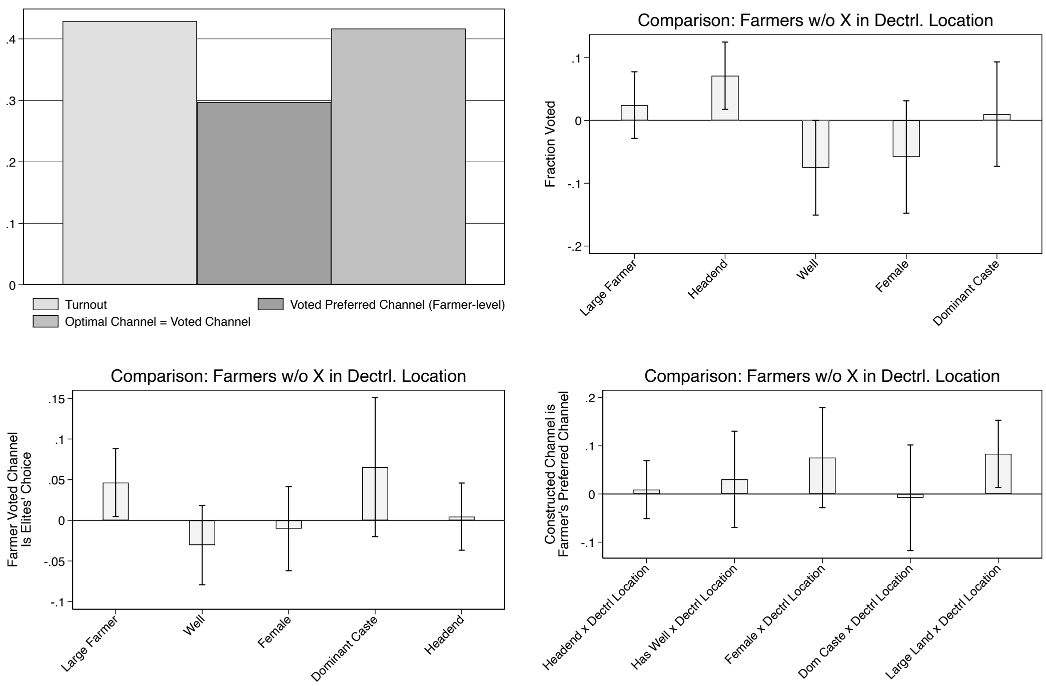
Notes: The images show the status of the constructed channel 3 years from the construction period.

Figure A3: Ballot Intervention



Notes: The images show the implementation of ballot or referendum intervention.

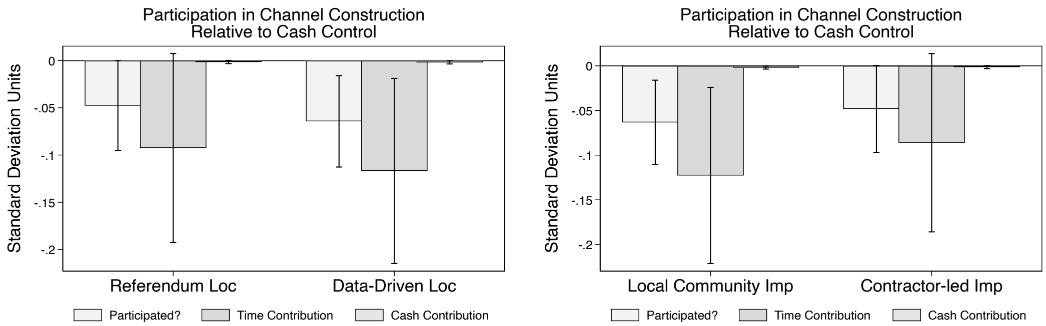
Figure A4: Referendum Outcomes: Using Administrative Data



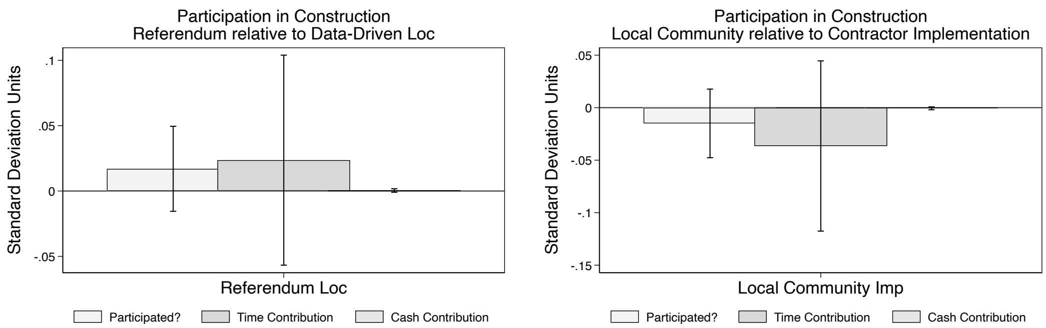
Notes: The figures examine turnout and outcomes of the referendum as well as differential patterns using administrative data on the referendum process by baseline farmer characteristics. These include differential turnout (top-right), whether voted channel aligns with elites' preference, and whether the constructed channel aligns with private preference by baseline characteristics. Baseline elicited value included as a control. Coefficients from the same regression.

Figure A5: Direct Costs of Participation

Panel A:

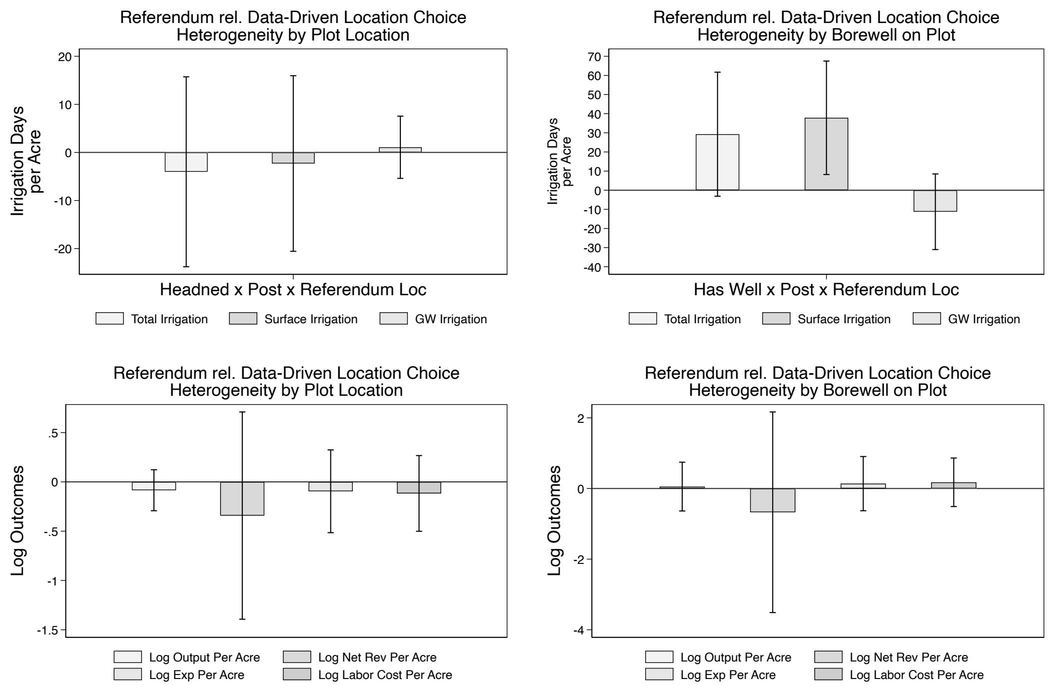


Panel B:



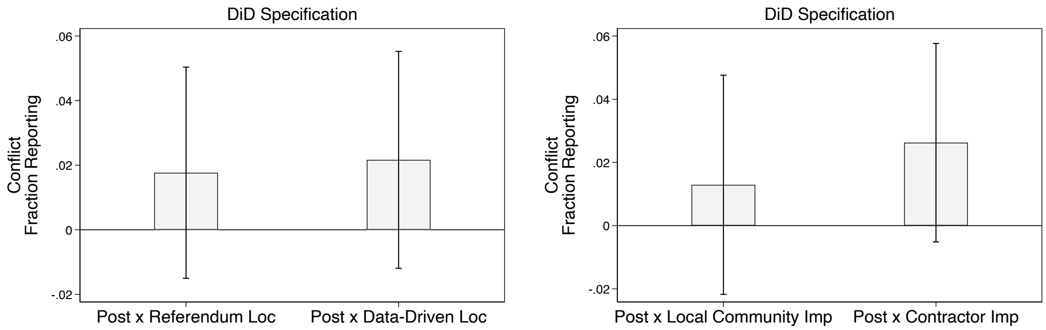
Notes: The figures document direct costs of participation in terms of whether or not the sample farmer participated in the channel construction, time and monetary contributions to the construction process. Panel A presents the average treatment effects relative to cash control where a channel was constructed in 20% of the villages. Panel B presents the difference between these participation outcomes by location choice and implementation groups, respectively. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure A6: Referendum Generates Regressive Distribution - By Other Prespecified Characteristics



Notes: The figures above present the heterogeneous treatment effect estimates on number of days of irrigation and production outcomes by plot location and by presence of borewell - groundwater resource. Each coefficient is the coefficient on the triple interacted term, interacting the baseline characteristic by the 2x2 program choice experimental status. The comparison group is the set of tailend/non-borewell plots in data-driven location choice villages. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Figure A7: Stated Conflict Outcomes



Notes: The figures above present the cost of participation in terms of reported conflict. We also measured the costs incurred by the local community in the construction and upkeep of any field irrigation channel in the study village/command area based on two survey questions - whether or not the respondent farmer participated in the construction and how much cash they contributed towards the construction, however, we find no differences in these outcomes across the experimental groups. The comparison/leave-out group is the budget-neutral control where some of the villages overcame the coordination frictions in constructing some channel on their own, without our facilitation. Standard errors are clustered by village/command area and the specifications include randomization strata fixed effects.

Table A1: Village-level Wage Effects

	(1) IHS(Wages) Agri Men	(2) IHS(Wages) Agri Women	(3) IHS(Wages) Non-Agri Men	(4) IHS(Wages) Non-Agri Women	(5) Wages Agri Men	(6) Wages Agri Women	(7) Wages Non-Agri Men	(8) Wages Non-Agri Women
Referendum Loc x Contractor Imp	-0.00136	-0.0489	-0.0105	0.00268	4.688	-27.54	-7.128	-4.443
	(0.0390)	(0.0444)	(0.0520)	(0.0598)	(30.34)	(23.27)	(40.96)	(29.02)
Referendum Loc x Local Community Imp	-0.0168	-0.0975**	0.0487	0.0256	-15.62	-43.75*	43.75	3.125
	(0.0389)	(0.0474)	(0.0546)	(0.0584)	(29.39)	(24.90)	(43.75)	(28.22)
Data-Driven Loc x Contractor Imp	-0.0549	-0.0529	-0.0179	-0.00625	-43.85	-14.47	-19.62	-11.44
	(0.0371)	(0.0482)	(0.0489)	(0.0572)	(28.14)	(26.39)	(38.24)	(27.11)
Data-Driven Loc x Local Community Imp	-0.0251	-0.0830*	-0.0172	-0.0129	-17.71	-37.08	-15.63	-12.08
	(0.0363)	(0.0428)	(0.0503)	(0.0591)	(26.83)	(23.90)	(39.66)	(28.42)
Obs	238	238	238	238	238	238	238	238
FE	District	District	District	District	District	District	District	District
Control Mean	717.71	448.96	779.17	477.08	717.71	448.96	779.17	477.08
Control SD	226.32	191.16	214.58	164.37	226.32	191.16	214.58	164.37
Ballot=No Ballot	.2835	.8883	.2878	.5318	.2117	.6147	.1874	.5304
Top Down=Bottom Up	.8015	.2905	.3848	.8306	.8851	.323	.3131	.8443

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: The table presents village-level agricultural wage regressions, where wage data was collected during the endline. Agricultural labor markets are segmented by gender. Male wage laborers are paid INR 700 per day in our context whereas female wage laborers are paid around INR 450 per day. The key distinction in the labor markets are that female labor is mainly used for weeding and harvesting whereas male labor is used during land preparation and fertilizer application. Standard errors are clustered by village and the specifications include randomization strata fixed effect.