[redacted]

[redacted]

1 Introduction

In parts of Northwest India water tables are being depleted at unsustainable rates where farmers grow water intensive crops such as rice (Fishman et al., 2011; Rodell et al., 2009; Russo et al., 2015; Shah et al., 2012). There are several factors contributing to the observed depletion. At the individual level, there is a lack of incentives to reduce water use in rice cultivation due to heavy subsidies and fixed fee pricing for electricity (Banerji et al., 2012; Meenakshi et al., 2011). Changing this pricing structure has had little effect on overall use in other states (Fishman et al., 2016; Meenakshi et al., 2011). This is particularly true in rice surplus states where rice production is subsidized and redistributed to other parts of India (Millet et al., 2013). In addition, erratic supply of electricity can lead to hoarding of water resources, as can an increase in the number of extreme heat days coming from climate change (Lobell et al., 2012; Taraz, 2018). Excess water compensates for the uncertainty of extremely hot days, when standing water evaporates quickly, and for the uncertain supply of electricity, which results in unpredictable irrigation. Individual farmers seem to heavily

discount future losses and even optimize their production process one season at a time conditional on the weather that they face. Another barrier to reducing water overuse occurs at the group level. Because wells draw water from a shared aquifer, a common resource, farmers' water use imposes a negative externality on other farmers' water availability, which individuals may not internalize, particularly if they do not share wells, as is often the case with ground-well irrigation (Kumar et al., 2011).

Given the above economic incentives, or lack thereof, training farmers in agricultural practices that conserve water might make little sense. Even without the many constraints listed above, adoption of new practices is often a slow process due to the risk, cost, and information asymmetries that exist in the agricultural context (Koundouri et al., 2016; Tsur et al., 2016).[1] More recently, however, behavioral approaches to encouraging adoption of green practices have been rather effective, even without any economic or price incentives behind them. Andors and Fels (2018) provide a systematic review of the types of behavioral "non-price" interventions that have succeeded in reducing resource use in the field, including commitment devices, goal setting, social comparison, and labeling. Key to each intervention's success is its ability to address well identified behavioral biases that can lead to excess resource use including: present bias, biased beliefs, costly information acquisition, bias toward concentration, exogenous inattention, and endogenous inattention (Alcott, 2016).

In this paper we ask whether and to what degree behavioral interventions that target both individual and group level norms can improve the adoption of a "green" practice, absent economic incentives. In addition, we provide selected farmers with a ground truth reference point - by metering the electricity use on their pumps - to test whether this information makes encouragement more salient. First, we introduce a well established agricultural practice called alternate wetting and drying (AWD), which reduces water use without reducing yields in a 2-arm randomized controlled trial (RCT). There are no immediate economic gains to adopting the practice as water pumping and electricity use are practically free and AWD does not, historically, increase yields. Rather, the practice appeals to the long term sustainability that is needed if farmers are to continue growing rice in the region. We introduce the practice in a way that can speak to some of the individual level biases mentioned above.[2] We establish well defined goals in terms of how much water to use each week given local conditions and provide regular reminders to address present bias; we provide free instruction via SMS and a free hotline for questions to make information freely available; and we emphasize continued monitoring of water pumping rather than a one time fix to support gradual rather than concentrated change. Next, we introduce a subtreatment arm in which we provide social comparison messages, where we compare a farmer's water pumping hours to other farmers' within the same farmer group.[3] In this way we can test whether addressing independent versus interdependent norms may drive adoption of a water saving practice.

[1] Foster and Rosenzweig (2010) provide a comprehensive overview.

[2] We focus less here on the use of labeling to address inattention given that we are introducing the adoption of a process rather than a product.

[3] Note that we do not estimate the effects of social comparison messaging alone, as other studies have done. It was essential that farmers were provided with the means and training to reduce their water use in safe, yet effective ways. Simply messaging them to do this could be detrimental to their livelihoods and food security without proper training.

Past studies focused on encouraging individuals to conserve resources have had promising results. Harding and Hsiaw (2014) demonstrate that setting reasonable goals for electricity consumption results in higher energy savings than for individuals with little to no goals or those with over zealous goals in Northern Illinois. Gilbert and Zivin (2014) show that following the receipt of an electricity bill, households reduce consumption by 0.6 to 1%. But encouraging conservation may not be enough if consumers have no reference point for acceptable levels of usage. For this reason, social comparisons, which highlight the interdependent nature of resource use, have become popular (Bicchieri and Dimatn, 2019; Meinzen-Dick et al., 2018). Allcott (2011) and Allcott and Kessler (2019) show that home energy use declines after home owners receive reports on their relative energy use in the US, and Aydin et al. (2018) demonstrates the same in Japan. Ferraro and Price (2013) show that home water use declines with social comparison treatments, particularly among high-consumption households. Datta et al. (2015) reports reductions of 3.7 and 5.6 percent in Belen, Costa Rica and Torres and Carlsson (2016) reports reductions of 5.8 percent in Antioquia, Columbia. While the latter studies investigate domestic consumption, Chabé-Ferret et al. (2019) is the only study that we are aware of that tests the effects of social comparisons on farmers' irrigation use in Southern France. They find small reductions of water use among the largest water users, but they also find an increase in use among farmers who previously did not use any water. Meinzen-Dick et al. (2018) also appeal to the interconnected aspect of water use in India, but rely more on communication between individuals to arrive at a collective choice rather than encouragement of each individual to conserve water in a decentralized decision making process. They use games to simulate crop choice and water use decisions. Participation in these games led to a greater number of communities adopting water registries and community governance of water.

In this study we focus on the degree to which non-price interventions can affect individual water use in a decentralized decision making context. We test whether a low-cost practice like AWD will be adopted by farmers if provided weekly reminders and instructions, or if farmers' beliefs about their own use require some updating - either via comparison to their fellow farmers or by simply knowing their own use through meters. Unlike past studies that test the effects of social comparison messages in primarily urban residential use settings, we present the effects in an agricultural household context, where water use simultaneously affects both household production and consumption. We do not touch upon the impacts of institutional changes in the provision of common resources as other studies do (Sekhri, 2011; Taraz and Sayre, Taraz and Sayre). As Agrawal (2001) documents, the large number of governance conditions that are required for successful management of common resources can be difficult to meet. We also control for other water savings practices such as land laser leveling (LLL), which has been shown to have larger water savings in India, but also can be quite costly (Lybbert et al., 2018).[4] First, we find that farmers' self-reports of their pumping hours track their metered values on average, showing that the metering may serve more as a reminder rather than a correction to farmers' beliefs regarding how much water they are using. Second, we find that there is a statistically significant reduction in pumping

[4]LLL adoption rates range from 25% to 46% depending on the treatment arm in our study, whereas starting adoption rates in Lybbert et al. (2018) were 1.5%

hours for the main treatment arm, as measured by both the meter readings and self-reported reading, in the order of two (self-reports) to five (meters) hours per acre per week, without any reductions to yields. The median farmer in this study's sample has 4 acres dedicated to rice cultivation, and pumps between 4 and 13 hours per acre per week (or between 16 and 52 hours each week) depending on the week and cultivation cycle. Third, when we couple AWD training with weekly social comparison messages (henceforth, AWD · messaging) we find a statistically significant effect for unmetered farmers. Thus, for unmetered farmers only, the treatment arm with social comparision messages is effective at reducing self-reported errors. The effect of AWD · messaging for metered farmers is variable. From these results, we can see that farmers do respond to the introduction of a water saving technique in the presence of reminders and goals, and that being metered appears to play a role in this impact. The social comparison messages did not have additional impacts on water savings for metered farmers, but they did seem to matter for the decline in self-reported hours of unmetered farmers. These results are promising, and ultimately could be used to complement other water conservation efforts like adequate land preparation and group decision making (Lybbert et al., 2018; Meinzen-Dick et al., 2018).

The paper is organized as follows. Section 2 lays out the context in which the trial took place. Section 3 describes the data collection. Section 4 summarizes the data, and Section 5 presents the econometric specification and results. Section 6 concludes.

2 Context and Study Area

The study took place in Haryana, India, across two districts: Kurukshetra and Kaithal –two northern districts where rice and wheat are the primary crops grown, where groundwater irrigation is prominent, and where ground water levels have been depleting (Aneja, 2017). The Central Groundwater Board of India monitors water quality and water depth at numerous observation wells across the country. Kaithal and Kurukshretra have 48 and 97 wells respectively. In 2015, the area of our study was classified as containing "deep" wells ranging from 20-40m in depth, with groundwater levels falling more than 4 m over the preceding decade, the highest level attributed by the study.[5]

Figure 2 shows the placement and proximity of the villages that we worked in. Haryana is one of the richer states in India, with per capita income growing faster than the national average since the mid 2000's. Unlike other parts of India, agriculture in the state is managed largely by relatively wealthier farmers who own their borewells and hire day laborers from neighboring states such as Bihar. In other states, such as Uttar Pradesh, for example, less wealthy farmers may rent another farmer's borewell at an hourly rate (Bubb et al., 2018). In 2013-14, the per capita income of Haryana was 133,427 INR, while the Indian average was 74,380 INR (Har, 2015). Haryana saw some of the highest growth next to Punjab during the Green Revolution largely due to the direct investments in infrastructure, information and incentives directed towards these states. Most notably, the government provided heavy

[5] Haryana State Central Ground Water Board, Ministry of Water Resources

electricity subsidies to these states. In the baseline, the total mean payment for electricity for the previous kharif season was 542 INR (approximately 8 USD), which was paid monthly at a flat rate per month via a state electricity bill. To place this in perspective, the average farmer in this study had 10 acres of land, 5 of which were devoted to rice, and reported approximately 200,000 INR (2900 USD) per acre per year in profits from cultivating rice.

Since the Green Revolution, farmers have expanded and grown their enterprises, and this would likely not have been possible without subsidized electricity (Bank, 2001). At this point, scaling back on production would not only set farmers back economically, but also India more broadly, as Haryana and Punjab are two of the largest providers of rice for the country. When farmers are asked what they would do if the government imposed unit pricing on agricultural electricity, most say they would not be able to maintain their current scale of production.

The local government does provide some training in sustainable agricultural practices. For example, drip irrigation training was available for farmers who purchased the pipes, via local agricultural universities, such as Haryana Agricultural University (HAU). However, diffusion appears to be slow and relies on farmers knowing of and attending trainings at a central location. Part of this may be due to the farmers' autonomy. Unlike small holder farmers with under an acre of land, farmers with their own irrigation tend to own more land, live farther from one another, and hire temporary day labor from poorer states. Thus, not only do farmers not observe each other's plots easily, they also do not interact as often as a small community might. This lack of interaction among farmers,[6] in conjunction with the high subsidization of electricity for agriculture, reduces incentives for water conservation.

3 Experimental Design

We worked with the Haryana Irrigation Research and Management Institute (HIRMI), a government supported water institute that provides local extension training to farmers. Figure 1 depicts the timeline of our study. The in-person baseline began in June 2015. We then conducted biweekly midline surveys via phone. The in-person endline took place in February 2016. The field experiment took place in 30 villages split across two districts, with 10 households per village, for a total of 300 initial participants.[7] We randomly selected 300 farmers to participate in the study in two stages. First, we obtained a list of the villages along the border of two districts, Kaithal and Kurukshetra, where over 50% of households grew rice, and which were located within an approximately 30 km radius of one another. The latter was meant to help ensure comparability across villages over time in terms of geography and weather. Next, we randomly chose 30 villages from this list, as seen in Figure 2, and from each selected village we randomly selected 10 rice growing households to participate in the study. Among the 300 participants, randomization occurred across two dimensions:

[6] For example, only 6% of farmers in our sample were aware of a farmer association in their village.

[7] In our initial power calculations we were powered to detect a 2.26 hour decline in water use allowing for cluster randomization `clustersampsi, detectabledifference mu1(8) sd1(6) m(30) k(10) rho(0.05)`. Mean and standard deviations were taken from our partners' data.

treatments and metering of pumps. We describe both below.

3.1 The Intervention

The two treatments were randomized at the village level, and the number of villages were evenly split across Control, Treatment 1 and Treatment 2. Treatment 1 included agricultural training in alternate wetting and drying (AWD) as well as weekly reminders of AWD practices from local scientists at the Haryana Agricultural Institute. Treatment 2 also received training in AWD and weekly reminders via SMS of AWD practices, but additionally received social comparison messages via SMS reminding farmers of their own pumping hours relative to other farmers' pumping hours (AWD messaging). The purpose of this was to provide an accurate estimate of average pumping hours.

We chose to train farmers in AWD (also known as controlled irrigation (CI)) with the consultation and advice of the International Rice Research Institute (IRRI). The practice reduces water requirements while maintaining or even increasing farmers' yields, and does not require additional labor or installation of equipment (Lampayan et al., 2014). Carrijo et al. (2017) provide a thorough meta analysis of 56 studies evaluating AWD with the overall finding that AWD reduces water use by 23% with no detriment to yields. Other studies such as Yang et al. (2017) find increases in yields with moderate AWD as it promotes stronger root growth of rice plants.

Beyond the potential benefits of AWD to yields, AWD can also lower costs to farmers who are facing declining water levels. On average, farmers in our sample reported a little over 0.61 m (2 ft) decrease in their water level per year since the time that their current pump was installed. Extracting water from deeper depths requires farmers to lower their pumps and sometimes purchase pumps with increased horsepower. Specifically, 83% percent of our sample had already lowered the pump in their wells, and 95% stated that they would need to significantly lower their pump in the next five years by an average of 17.4 m (57 ft). This could cost on the order of 25,000 INR (360 USD) based on the average response from our surveys. Furthermore, new pumps with horsepower over 20 (the average farmer had a horsepower of 19 (median of 20) can range from 500 to 1500 USD.

Many farmers reported in the baseline that they were already trying to mitigate these costs. Ninety-six percent said that they anticipate water shortages in the near future. About 43% of our respondents reported attempting to improve their water conservation on their own by manually controlling their pump use (48%), by changing farming techniques (25%), field leveling (100%), and rain water harvesting (47%). In addition to the increased cost associated with pumping water from deeper levels, depleted aquifers typically contain older water that is not recharged, leading to unsustainable groundwater depletion. Continued over-extraction can lead to aquifer salinization, which is poor for crop yields (Dasgupta et al., 2018; Greene et al., 2016; Okorogbona et al., 2015).

The first introduction to AWD training took place after the baseline survey was completed in farmers' respective villages. The next training took place at HIRMI where HAU research

scientists discussed AWD's potential benefits for rice and to reducing long-term watering costs in more depth. The researchers then reviewed the steps associated with AWD. IRRI provided the leaflets and photos for AWD training (see Figures 3 and 4), which we translated into Hindi.[8] In addition, IRRI provided us with the specifications to manufacture the pani pipes in Haryana. The pani pipes enabled farmers to monitor their water levels.

The primary AWD steps are relatively succinct. Two weeks after transplanting rice saplings from the nursery bed to the main field, the farmer should flood his field to a depth of 5 cm, then allow it to drain approximately 15 cm below the soil surface. Once dried, the farmer should re-flood his field. This process continues until two weeks before flowering. At the flowering stage the farmer should continually flood his field. In this study, farmers received weekly text messages with the suggested requirements for AWD at each stage - namely whether to keep fields flooded, and at what level, or whether to let fields dry. The weekly reminders via SMS were created with the guidance of HAU researchers who instructed farmers on AWD in the training session and who were located in Kurukshetra district and could monitor local conditions and rainfall. These same researchers also provided an additional service by taking calls from any farmer in the training who had questions about AWD procedures for the duration of the study.

3.2 Metering

To quantify the effects of the treatments, we need estimates of water pumping use at an individual level. Past studies investigating water use in the developed country context have often relied on precise metering devices (Allcott, 2011; Ayres et al., 2013; Chabé-Ferret et al., 2019; Ferraro and Price, 2013), where the infrastructure is already in place. In the developing country context, studies measuring water use typically rely on national surveys that measure groundwater depths at the well level on an infrequent basis, or rely on satellite-based estimates of groundwater depth. For example, Sekhri (2011, 2014) uses the Minor Irrigation Census (MI Census) conducted by the government of India every five years. Rodell et al. (2009) use water storage-change observations taken from the NASA Gravity Recovery and Climate Experiment satellites to estimate total depletion rates. Fishman et al. (2015) simulate aggregate depletion rates based on parameter estimates taken from various publicly available data sources. Conversely, our study is interested in quantifying individual level effects, and aggregate measurements would not allow us to detect changes in individual use.

One possibility is to only elicit individual self-reported use via surveys. This can be problematic for two reasons. First, even if farmers know how much water they use in a week, their self-reports may suffer from recall biases. The literature on biases arising from self-reporting is well established.[9] Recent studies detail the evidence around recall biases in the developing country context. Bell et al. (2019) show that respondents' recall is worse for consumption and experience outcomes than for farm activities (but this did not include irrigation). Beegle et al. (2012) show that recall bias is lower for financially taxing events, such

[8] See: Saving Water with Alternate Wetting Drying (AWD)

[9] Evans and Leighton (1995) and Sudman and Bradburn (1974) provide an overview.

as purchasing costly inputs, in Malawi, Kenya and Rwanda. This would suggest that recall bias could be potentially high for water use, which costs farmers very little. de Nicola and Gineé (2014) find that recall bias increases over longer lengths of time in recalling income and asset purchases in India. Godlonton et al. (2018) find significant anchoring bias occurs in agricultural surveys for both subjective and objective measures in Central America.[10] Second, eliciting self-reported water use can be challenging because farmers may legitimately not know how much water they pumped in a week. A farmer can set his or her pump to run automatically, so that it turns on when there is electricity. If the farmer does not visit his or her pump on a daily basis then he or she may not be aware of the times it was running.

Given the extensive literature studying biases in self-reported survey outcomes, our study aims to estimate how much pumping hours differ between metered and self-reported reports. Fishman et al. (2016) is another study that investigates this difference. In their study, metering is voluntary and directly incentivized - farmers receive compensation for saving electricity above a benchmark. Overall, they find that farmers strategically self report their electricity use - often under reporting their use in relation to the physical metered reports. The role that metering plays in this paper serves a different purpose than in Fishman et al. (2016) in that being offered a meter is randomly assigned, and we are interested in how that assignment affects self-reported water use. One difficulty with this design was that for a handful of farmers we were unable to successfully install a meter, and for others the meter failed to work for the duration of the study. We address some of the imbalances this may have caused between treated and control groups in Section 4.

An additional merit of this study is to be able to compare farmers' self-reported versus metered pump use on an hourly basis. As mentioned above, there are well documented recall biases found in individual level agricultural survey data. Without prior studies on water use in this region, it would be difficult to know if self-reported water use is accurate. Note that the ideal outcome variable for capturing actual water use would be volumetric measurements of the water pumped by each farmer. However, attainment of this information would require an inline flow meter that measures cumulative flow and alteration of the irrigation pipes for installation at every household, costing hundreds of dollars each. Further, we were interested in an outcome variable that would enable us to compare what farmers could reasonably self-report versus a ground truth measurement. Had we captured volumetric measurements, it would have required several assumptions regarding pump power, pump efficiency, and actual water depth being pumped and would have been harder to compare to farmers' self-reported pumping hours. Thus, we decided to capture the number of hours that farmers self-reported from their own recollection, the pump's horse power and age, as well as the farmers' reports of a meter reading, described below.

We randomly selected 200 households (approximately 66 households per treatment arm) from the list of 300 households to have their pumps metered, conditional on their consent, and no household declined. The option to have one's pump metered was introduced alongside the aim of AWD - monitoring water use - and farmers were informed that monitoring was the

[10] Godlonton et al. (2018) define anchoring bias as "when individuals use some easily-observed prior or recently-provided information to guide them in estimating a value under uncertainty."

sole purpose of the meter. For example, the notion that the meters might be used to then introduce unit pricing did not come up. Further, the information was delivered by trusted regular local extension agents from HIRMI and HAU. Of the 200 who were offered meters, only 154 meters could be successfully installed.[11,12] The final breakdown of metered versus not metered household across all treatment arms is shown at the bottom of Table 2. In Section 4 we discuss the implications of being successfully metered and the extent to which the resulting number of metered households affected balance across treatment and control groups.

We hired a team of three local electricians to install the electricity meters, which attached at the circuit breakers of the water pumps. We metered the pump on the main rice plot, which was the largest contiguous rice plot that the farmer owned and about which the farmer would be answering questions. If the pump served more than one plot, we controlled for this by capturing the percent acreage that the pump served for the main rice plot for which we surveyed. Less than 4% of our sample had a pump that served more than one rice plot. The electricity powering pumps were enclosed in small brick structures, and served only the pump at that location. The watering pump was the only device connected to the water meter circuit; hence the meter only recorded activity when the pump was running. The pumps were all 440 volt industrial powered units that connected to the circuit. If the electricity was turned on, but the pump was not running, the meter was not recording electricity use, which we verified at the time of installation.

3.3 Text Messaging and Surveys

Text messaging was used to deliver information for the two treatment groups, which included reminders for following the AWD technique as well as reporting hourly pumping. In addition, all households, including those in the control group, were sent text messages to remind them of the weekly phone survey, which occurred on Mondays. Thus, all households were receiving some form of text messages. Control villages were reminded of the weekly survey; the AWD villages were reminded of the weekly survey and AWD directions; and AWD · messaging villages were reminded of the survey, the AWD direction for the week and their hourly pumping use relative to the other participating farmers in their village. The text message for the weekly survey read as: "We will call you tomorrow for the weekly survey." Metered households were also reminded to check their meter in this message, as we would

[11] Due to budget constraints, we did not have the resources to meter all 300 households.

[12] Many electrical configurations in the fields were not compliant with code, especially regarding the electrical ground. Electrical grounding is necessary to provide a constant reference for accurate voltage measurements. In the most rudimentary configurations, the grounding wire is directly connected to the earth; however, depending on the soil moisture, these grounding configurations require saturation, introducing uncertainty as to whether or not the meter would measure usages. Due to the ad hoc nature of the electrical wiring, even with proper meter installation, the meters would not log all use, but rather only record the use when the grounding wire also received adequate water. Finding additional households to survey and meter was also not feasible, both because labor supply was particularly limited during the month of May when temperatures reached highs of 44 degrees Celsius (111 Fahrenheit), and because the study was nearing the start of paddy transplantation, before which training had to be completed. For these reasons the final count for households where we could successfully install functional meters was 154.

ask for its value during the phone survey.

The AWD treatment group received a weekly message with the following format: "Continue with AWD. Total expected rainfall for the next 5 days is XX (YY) mm in Kaithal (Kurukshetra) district," followed by what was necessary that week depending on if it was flooding season (dry field to 6th pani pipe hole after standing water reaches 50 mm), flowering (maintain 5 mm standing water) or harvesting (dry field). The AWD·messaging treatment received the latter AWD messages as well as a message that read, "Your pumping hours are XX and your group's average pumping hours are XXX. Conserve water collectively." The average was taken at the farmer group level and came from the previous week's survey information. We chose to use the farmer group in each village as the unit of measure, as farmer associations are one of, if not the most, pertinent reference network for agricultural knowledge in this region. Further, HIRMI and HAU - the two partners that helped implement this study, provide government supported training via the well established farmer group infrastructure. Farmers typically make decisions and are informed of local knowledge and training through this network.

Surveying took place in two ways: two in-person surveys for the baseline and endline, and then eight weekly mid-line phone surveys.[13] All farmers were surveyed based on a representative rice plot throughout the season. The main outcome variables of interest are a) self-reported pumping hours and b) the meter value attached to the farmers' pump. Note that farmers reported both measures voluntarily by phone. Seventy-five percent of the farmers who were metered consistently picked up their phones.

For self-reported pumping hours, we asked the farmer to recall how many hours he pumped each day for the previous week. For the meter reading, we asked the farmer to simply read off the number on the meter. The farmers were not explicitly told what the number represented, which was a sequence of minutes, seconds, and milliseconds. We recognize that there could also be reporting errors with the meter reading that may raise concern. We address this in two ways. First, the meter reading is cumulative; therefore, the farmer should be reporting increasingly higher numbers each week. In 20 instances, we had meter readings that decreased from one week to the next. We exclude these observations. Second, we did a final in-person check of the meter readings to verify that the last value the farmer reported over the phone and the reading we took in person via enumerators were in a reasonable distance from one another, which they all were.

4 Data and Descriptive Statistics

The average farmer in our sample is a 46-year-old married male with about 10 acres of land (half of which was allocated to cultivating rice) with access to groundwater irrigation.[14] These are medium sized farms operated by the male head of the household. Women in these

[13] We were unable to complete surveys in week 3 due to cell tower outages. Surveys in the control group did not start until week 2 due to miscommunication with enumerators making calls.

[14] Because all the main respondents in our study were male, we refer to farmers using he, him, his pronouns.

households may participate in the field work but are not the decision makers. We do not have a gendered breakdown of labor, but we can observe that for certain tasks (nursery bed and land preparation) the total members of the household equaled the total household members utilized for the task, thus, women were likely involved in the task. For other tasks, such as transplantation, nearly 80% hired outside labor in addition to using some family members. Other studies in this area of India find similarly low participation rates of women in agriculture (Mittal and Hariharan, 2018).

To ensure that our randomization of treatment and control groups was performed successfully we test that observable covariates are balanced across all three groups: Control, Treatment 1 and Treatment 2, before the start of the intervention using baseline data. Table 2 reports the results. Our sample appears balanced across treatment and control groups for all observable variables except for respondent's phone ownership. Farmers who did not own their own phone could still participate if they were able to use a friend's or household member's phone. In addition to investigating the balance of the full sample, we also test for balance between farmers who were and were not metered. In Table 3 we see some imbalances despite randomization, including for age, and marital status. We control for these covariates throughout.

We do not capture characteristics of each farmer's field, such as geological and soil types, which would have been prohibitively costly to sample. Thus, we cannot speak to whether treatment versus control villages are systematically different along these dimensions. That said, according to maps created by the Haryana Space Application Center (HARSAC), Kurukshetra and Kaithal contain similar geological characteristics and similar soil types.[15] According to the Central Research Institute for Dryland Agriculture the soil in these regions is loam and sandy loam.[16,17] These soils typically have moderate infiltration rates and maps of the district are not sufficiently accurate to control for changes in soil characteristics on a plot level. Given the similarities in expected infiltration rates, we normalize all self-reported and metered irrigation hours by acres of rice irrigated by the well in question.[18] Physiographically, the northern part of Kaithal is similar to Kurukshetra consisting of upland plains. The villages are all located within a 32 km radius, which also limits the potential geographic variation between the villages. We do control for total weekly rainfall at the village level in our estimations, which helps control for some physiological variation across villages. We use weekly rainfall estimates from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), which we plot in Figure 5. The year 2015 was a relatively poor year with regards to rainfall, particularly from week four and onward.

[15] Geology Map of Haryana.

[16] The soil is 75% loamy and 19% sandy loam in Kaithal, and 46% loamy and 38% sandy loam in Kurukshetra.

[17] Gyawali et al. (2016) also characterize the soil of Kaithal and Kurukshetra as varying from loamy sand to clay loam.

[18] It is standard convention to standardize applied volume of irrigation water by acres irrigated (gallons per acre). We used the standard equation $Q*t = d*A$ to justify our normalization of all hours (t) by acres irrigated (A), assuming little variation in applied flow rates (Q), thereby serving as an approximate depth of irrigation water applied (d).

5 Results

Our summary statistics from the endline indicate that farmer adoption of the pani pipes within the treatment groups was 46%, where 20% said they checked their water levels daily, 25% checked 1-2 times per week, 32% checked weekly, 20% checked biweekly, and 3% did not reply. While adoption was not 100%, it is high in comparison to adoption levels seen in first time use even for agricultural techniques that are yield enhancing.[19] Further, even if farmers in the treatment groups did not use the pani pipe, it did not preclude farmers from implementing the recommendations they received on a weekly basis. Many farmers use other techniques for measuring water levels such as a their fingers or a ruler.

We estimate the intention-to-treat for Treat 1, AWD alone, and Treat 2, AWD · messaging, for farmer i in village j on hourly pumping. In addition to the main treatments, in some specifications we also include a variable for whether a farmer's metering is above or below the village average in pumping hours in the previous week, *L_above*. We interact this variable with the second treatment arm, AWD · messaging, since low-use farmers are essentially receiving a different treatment than high use farmers, as we need to account for this differential effect. We control for any variables for which we found imbalances on (marital status, owning ones own phone, age, drinking source) in Table 2 and Table 3, as well as village level total weekly precipitation throughout. We also allow for week fixed effects, w_t, and farmer specific random effect, γ, for each i farmer.[20] Equation 1 reflects our specification:

$$\begin{aligned}\mathbf{Hours_{ijt}} = \alpha + \beta_1 \mathbf{AWD}_j + \beta_2 \mathbf{AWD} \cdot \mathbf{messaging}_j + \beta_3 \mathbf{L_above}_{ij} \\ + \beta_4 \mathbf{AWD} \cdot \mathbf{messaging} \cdot \mathbf{L_above}_{ij} \\ + w_t + \mathbf{X}'_{ij}\boldsymbol{\theta} + (\gamma_i + \epsilon_{ijt})\end{aligned} \tag{1}$$

To provide some context of the outcome variable's magnitude, an additional hour of water can deliver approximately 21,460 gallons/hr (84,944 L/hr).[21] Translating this figure into aquifer replenishment rates would require a hydrologic model, which is beyond the scope of this paper. We provide it to illustrate that a change in farmer pumping behavior, even in the order of an hour, has an economically significant impact.

Before turning to our model estimates, we plot weekly pumping hours according to reports from the meters versus self-reports (Figure 6), where the four groups that we plot are shown

[19] Macours (2019) provides a summary of adoption rates for yield enhancing, variance reducing and labor saving techniques.

[20] Since treatment is randomly assigned, and we are interested in the between effect, between individuals i across different treatments that do not vary with time, a random effects models allows for different intercepts for each individual in this context. Furthermore, because treatment assignment is random, person specific intercepts will not be correlated with treatment assignment (Bell et al., 2019).

[21] We can convert hours of pumping to approximate pumping flow rates and volumes using a basic horsepower (hp) calculation (flow rate (gallons/min) = $\frac{\text{hp}X\text{pump efficiency}X3960}{\text{depth}}$). Assuming a 20 hp pump, efficiency of 70%, a water depth of 145 ft (44 m) and an additional 10 ft for the pump's placement, as well as no flow restrictions due to pipe friction or diameter, the pump could theoretically deliver approximately 20x3960x.7/155 = 374 GPM (gal per min) * 60 min/hour = 21,460 gal/hr.

in Table 1. In Figure 6, we compare the self-reports, which include values from farmers who were and were not metered (I and II in Table 1), to the metered reports (III in Table 1). Overall, we can see that average self-reports and reports from meters largely follow the same u-shaped pattern (week-by-week means are not statistically different except for week 2, which we discuss later). This u-shape concurs with the general irrigation patterns for rice, which begins with active flooding that dissipates over time, and then active flooding during the flowering period, right before harvest. We are also interested in how self reports differ depending on whether farmers' pumps are metered or not (I compared to II in Table 1). In this case, we are not comparing the same farmers. This is an interesting point of comparison because farmers without a meter lack the reference point that metered farmers have and their self reports may differ. In Figure 8, we plot the two series and see some discrepancies, where usage is higher for metered farmers in the AWD·messaging arm in weeks 5 and 6 (pvalue = 0.03 and 0.07 respectively).

The latter figures demonstrate that *within* each treatment arm the comparison of metered and self reported values largely concur with one another. Now we turn to how the treatments affect both measurements by comparing outcomes *across* treatment arms. Table 4 presents the estimates from Equation 1, which we estimate separately for metered outcomes and self-reported outcomes. Column 1-2 are the estimates of the treatments' effects on metered hours per acre, that is, comparing the metered reading series in Panel C and D in Figure 6 to Panel A. Columns 3-8 are the estimates of the treatments' effects on self-reported hours per acre. Columns 5-6 are self-reported pumping hours reported by farmers who did not have a meter installed at their pump (the self-reported series of unmetered farmers in Figure 8), and Columns 7-8 are self-reported pumping hours reported by farmers who did have a meter installed at their pump (the self-reported series of metered farmers in Figure 8). All columns include the aforementioned controls: age, owning one's own phone, marital status, private drinking source and total weekly village-level precipitation. In Column 1, we can see that AWD alone had a negative effect on pumping hours of about 3.7 - 5.1 hours per week per acre. Note that weekly pumping per acre can vary from 4 to 13 hours depending on the week, and so this reduction averages over both smaller and larger changes over the course of several weeks. In terms of self-reported hours, we see a 1.5-1.8 hour decline in hours per week per acre caused by the AWD training, and this is driven by farmers with meters, as shown in Columns 7 and 8. Farmers without meters in the AWD arm do not report a decline in water use. Since the direction and magnitude of the effects from AWD training are similar for metered and self-reported hours, it is likely safe to conclude that AWD training did indeed help farmers conserve water. However, the sample size for metered farmers is smaller than for what we had initially powered our study, and so we cautiously interpret the effect sizes for the metered sample, which are, therefore, likely to be overstated given wider tails over the distribution (Gelman and Carlin, 2014; Vasilaky and Brock, 2020).[22] That being said, the difference between the effect sizes for metered farmers in Column 1(2) and 3(4) are not statistically significantly different from one another (Chi squared pvalue = 0.48(0.25)). Further, given the smaller sample size and number of village clusters we also

[22]The lower the power of a test for a given variance, the closer that the null, H_0, and the alternative, H_1, hypothesis will be. A more extreme point estimate is, therefore, needed in order to reject H_0 in favor of H_1 in a low powered study (Vasilaky and Brock, 2020)[pg 9]

wild bootstrap our standard errors, which is reported in Appendix A.[23] The patterns in our results are largely the same. Although, as expected, the weakly significant effect in Column 1 is no longer statistically significant.

In terms of cost savings the majority of farmers reported that their pump was fueled by public electricity rather than a private generator (only 5 farmers reported using a generator to power their pump). Our estimates of costs for pumping are crude, since we solicited the lump sum cost for the entire season for one pump. The average cost per season per pump was 540 INR ($\sim$ 7 USD) (maximum of 5000 INR per season ($\sim$ 70 USD)). Factoring in the final cumulative sum of hours pumped according to metered farmers, the cost per hour per metered farmer was a mere 1.5 INR. Given such low, albeit crude, estimates of costs, it is difficult to imagine that cost savings on electricity usage provided any incentive to conserve water.

Overall, it appears that AWD and the structured reminders targeting individual water use had a significant impact for metered farmers, and this is reflected in both metered and self-reported outcomes. One potential mechanism for this effect is that the meter helped anchor farmers' use, allowing each farmer to become their own point of comparison over time as they were provided with AWD instructions for wetting and drying. Figure 7 provides some suggestive evidence of this. It depicts only the farmers who were metered in the study and plots both their self reported and metered values in the same graph. We can see that largely the two values closely match one another, except for the initial week. We can see that there is a significant difference between metered and self reports in the first week that both measures were reported - week 2. The metered readings far exceed the self-reported readings in week 2, but, interestingly, by week 4 the two reports coincide with one another, suggesting that the metering potentially anchored farmers' perceptions of their use.

Turning to the effect of the AWD·messaging arm, we see no impact on metered pumping hours in Columns 1 and 2. Recall that this treatment arm still included training, thus the additional messaging seemed to have dampened the effects of the AWD encouragement.[24] In terms of the effect on self-reported hours, we do see a 1.4-2.2 hour decline for Treat 2 in Column 3 and 4, and now unmetered farmers are contributing to this reduction in self-reported water use. In fact, unmetered farmers are responding in their self-reported hours to Treat 2 - the training with social comparison messages - but not to the training alone. One possible reason for this may be that social comparisons are effective when farmers do not have a reference point for their own metered values making the group average a more salient reference point. Another possibility is that the social comparison message affected unmetered and metered farmers differently, where it decreased use for unmetered farmers, but actually increased use for metered farmers. To support this, Column 8 shows that once

[23] Note we wildbootstrap the errors using the `clustse reg, method(wild)` command in Stata, which does not allow for random effects. For this reasons our point estimates also change given the loss in efficiency of estimators.

[24] We chose a design in which social comparisons were only added to the main training arm. Simply encouraging farms to reduce their water use, as studies have done with residential use (Datta et al., 2015; Ferraro and Price, 2013; Torres and Carlsson, 2016), was not feasible, nor ethical, as farmers may not have had the techniques and skills to do so without compromising their livelihoods and food security.

we control for L_above -whether a farmer is above the village farmer group's average use for the previous week- metered farmers' self reports reflect a reduction in water use in response to Treat 2 (but it is still not larger than the effect of AWD alone).[25] However, the positive sign on the interaction term of $L.above \cdot T2$ suggests that social comparison may be encouraging frugal water users to continue conserving, and bigger water users to continue using. That is, the comparison, at least among metered users, may have helped reinforce rather than change existing behaviors. Bicchieri and Dimatn (2019) speak to the possibility of this occurring with social comparisons and that "publicizing any behavior could lead to the recipient believing that that behavior is approved of." Indeed, with the social comparisons, farmers were not explicitly told that they must reduce their use or that they were wasteful if they were above the farmer's group average. It may very well be that they justified their use in the current week given that the reports merely stated their reference group's average from the previous week.

Our results show, first, that metered and self-reported hours largely track one another *within* treatment arms across most weeks as depicted in Figure 6. This is useful information as collecting self-reported values is less expensive than metered values, and possibly a way to enhance our knowledge of water use patterns. We do find statistically significant discrepancies *between* treatment arms. First, we see that for farmers who were metered, the AWD training and messaging had a statistically significant downward effect on water use, and this finding is true if we use the readings taken from metered pumps (III in Table 1), or if we use the self-reported readings of metered farmers (I in Table), and the effects are not statistically significantly different. Next, we find that layering on a social comparison message has a less straight forward effect. When farmers are not provided a meter to monitor their use, social comparisons appear to encourage a self-reported reduction in water use. However, when farmers are metered the social comparison can reinforce rather than change existing norms. In past studies, social comparisons have been successful in helping to reduce residential water consumption, but they seem less well suited to the more complex nature of an agricultural household in which water use is tied to the household's livelihood. From a policy perspective, our results suggest that the cleanest way to help farmers conserve water is by providing them with guided encouragement at an individual level while also making metering available and observable.

5.1 Consumption, Yields and Cooperation

In this section we explore how farmers' choices in terms of pumping hours translated into input and output outcomes, as well as feelings of cooperation and trust in other village members. We use the endline survey outcomes recorded at the end of the study to estimate these effects cross-sectionally. In particular, we want to ensure that AWD did not have any detrimental effects on yields, as this was one of the key selling points of the practice. The fact that AWD did not promise yield gains also helps us gauge farmers' true desires to regulate their water use.

[25] This result differs from the results in Column 1, which show no effect of Treat 2. It is possible that the metered sample in Columns 1 and 2 may not be powered to detect smaller effect sizes and this may contribute to the differences between these two results.

Regarding inputs, including labor, fertilizer and pesticide, we see no changes due to the treatments in Table 5. With respect to output, Table 6 shows that farmer-reported yields (kilograms per acre) were not affected, but also imprecisely measured. There is suggestive evidence that yields increased for unmetered farmers but decreased for metered farmers, who we know reduced their water pumping hours. Rice consumption, reported as the amount of rice consumed over the last 30 days, increased significantly for AWD · messaging. This self-reported effect is also driven by the unmetered group in Panel B, whose consumption per household increased by as much as 3 kgs per month. Together, these results provide suggestive evidence, consistent with research cited earlier, that AWD training enabled farmers to reduce their water use, and did not reduce output or consumption of rice (Carrijo et al., 2017).

Table 7 estimates the effects of the treatments on whether farmers believed their community was cooperative, and whether they could trust that other farmers in the community would do what is best for the community and not only for themselves. Column 1 (Cooperate) shows if farmers believed other farmers in their community were cooperative with one another. Column 2 (Trust) shows if farmers believed that other farmers in their community would do what is best for the whole community. We can see that beliefs in cooperation increased by 2.6 - 4.4 % for both treatment arms, but beliefs in trusting others did not change. This change does not appear to depend on whether the farmer's pump was metered or not.

6 Conclusions

In this paper we compare self-reported versus metered measures of water pumping hours in the presence of a randomized intervention in which farmers are trained in Alternate Wetting and Drying (AWD) and receive social comparison messages of their village peers' average water use. Unlike past studies that test the effects of social comparison messages in primarily residential use settings, we present the effects in an agricultural household context in rural India, where water use affects households' consumption and production decisions.

There are several key takeaways. Overall, self-reported and metered hours largely track one another from week to week. Farmers with meters on their pumps report hourly usage that is similar and not statistically different from recall versus meters. Furthermore, self-reported values between farmers who were and were not metered are also similar on average, and not statistically different within villages. We find that encouraging farmers to conserve water at an individual level is effective if a farmer's water pump is metered. Metering may provide the reference point needed for farmers to follow the AWD instructions. This is a useful finding, especially when the majority of encouragement designs regarding agricultural technologies are centered around yield enhancement rather than conservation. Next, we find that adding on social comparison messages has a varied effect. For farmers who are metered, the message may have had adverse effects where it increased use by high users and decreased use by low users, and dampened the overall effect of the AWD encouragement. For farmers without meters, social comparison messages appear to have aided in the effectiveness of the AWD

encouragement, possibly because it provided a reference point that unmetered users did not have.

The current study provides a starting point for understanding in what ways non-price interventions can encourage water conservation, especially when price-based interventions have not had considerable success. Our results suggest that in this agricultural household context encouraging conservation along with observable monitoring is the most direct approach for affecting water use in the region. Our research does have several limitations, and calls for additional study. First, our sample size, particularly for metered farmers was small and, therefore, the statistically significant effect of AWD is likely an overstated effect size, and we are underpowered to detect effects from the AWD·messaging arm. Next, while the effects we estimate are an intent-to-treat (farmers were not required to adopt the pani pipe nor did they have to implement the AWD suggestions), response to our phone surveys was entirely voluntary each week. Therefore, the data and responses reflect that self selection. Future work with larger samples and resources could incorporate smart meters that push logged values to the cloud, so that researchers can view the meter readings in real time. Smart meters could also inform researchers if devices are failing. Weekly in-person interviews with a larger sample could help with response attrition. Finally, more nuanced experiments are needed to understand the underlying effects of combining feedback on one's own use and feedback on peers' use. There is literature that supports the overall efficacy of providing feedback in energy conservation (Karlin et al., 2015; Wok et al., 2015), but less is known about the interaction of feedback coming from multiple sources. Our results call for greater understanding of these effects, especially in the less studied, but urgently needed, context of agricultural use.

References

(2015). White paper on state finances, finance department http://www.finhry.gov.in/Portals/0/whitepaper-finance.pdf. Technical report, Government of Haryana.

Agrawal, A. (2001). Common Property Institutions and Sustainable Governance of Resources. *World Development 29*, 1649–1672.

Alcott, H. (2016). Paternalism and energy efficiency: an overview. *Annu. Rev. Econ.*, 145–176.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics 95*(9), 1082–1095.

Allcott, H. and J. Kessler (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics 11*, 236276.

Andors, M. and K. Fels (2018). Behavioral Economics and Energy Conservation A Systematic Review of Non-price Interventions and Their Causal Effects. *Ecological Economics*, 178–210.

Aneja, R. (2017). Ground Water Level in Haryana: A Challenge for Sustainability. *International Journal of Research and Analytical Reviews 43*(3).

Aryal, J. P., M. B. Mehrotra, M. L. Jat, and H. S. Sidhu (2015). Impacts of laser land leveling in ricewheat systems of the northwestern indo-gangetic plains of India. *Food Security 7*, 725738.

Aydin, E., D. Brounen, and N. Kok (2018). Information provision and energy consumption: Evidence from a field experiment. *Energy Economics 71*, 403–410.

Ayres, I., S. Raseman, and A. Shih (2013). Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage. *The Journal of Law, Economics, and Organization 29*, 992–1022.

Banerji, A., J. Meenakshi, and G. Khanna (2012). Social contracts, markets and efficiency: Groundwater irrigation in North India. *Journal of Development Economics 98*(2), 228–237.

Bank, W. (2001). India - Power supply to agriculture (Vol. 2): Haryana case study (English). *World Bank Report 22171*.

Beegle, K., C. Calogero, and K. Himelein (2012). Reliability of recall in agricultural data. *Journal of Development Economics 98*(1), 3441.

Bell, A., M. Fairbrother2, and K. Jones (2019). Fixed and random effects models: making an informed choice. *Quality Quantity 53*(5), 10511074.

Bicchieri, C. and E. Dimatn (2019). Nudging with care: the risks and benefits of social information. *Public Choice*.

Bubb, R., S. Kaur, and S. Mullainathan (2018). The Limits of Neighborly Exchange. *Working Paper*.

Carrijo, D. R., M. E. Lundy, and B. A. Linquist (2017). Field Crops Research Rice yields and water use under alternate wetting and drying irrigation: A meta-analysis. *Field Crops Research 203*, 173–180.

Chabé-Ferret, S., P. Le Cont, A. Reynaud, J. Subervie, and D. Lepercq (2019). Can We Nudge Farmers Into Saving Water? Evidence from a randomized experiment. *European Review of Agricultural Economics 40*(3), 393–416.

Dasgupta, S., M. M. Hossain, M. Huq, and D. Wheeler (2018). Climate change, salinization and high-yield rice production in coastal bangladesh. *Agricultural and Resource Economics Review 47*(1), 6689.

Datta, S., J. J. Miranda, L. Zoratto, O. Calvo-Gonzlez, M. Darling, and K. Lorenzana (2015). A Behavioral Approach to Water Conservation Evidence from Costa Rica. *World Bank Research Working Paper* (WP7283).

de Nicola, F. and X. Gineé (2014). How accurate are recall data? Evidence from coastal India. *Journal of Development Economics 106*, 5265.

Evans, D. and L. Leighton (1995). Retrospective Bias in the Displaced Worker Surveys. *The Journal of Human Resources 30*(2), 386–396.

Ferraro, P. and M. K. Price (2013). Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *The Review of Economics and Statistics 1*(95), 64–73.

Fishman, R., N. Devineni, and S. Raman (2015). Can improved agricultural water use efficiency save India's groundwater? *Environmental Research Letters 10*.

Fishman, R., U. Lall, V. Modi, and N. Parekh (2016). Can Electricity Pricing Save India's Groundwater? Evidence from Gujarat. *Journal of the Association of Environmental and Resource Economists 3*(4), 819–855.

Fishman, R. M., T. Siegfried, P. Raj, V. Modi, and U. Lall (2011). Over-extraction from shallow bedrock versus deep alluvial aquifers: Reliability versus sustainability considerations for India's groundwater irrigation. *Water Resources Research 47*, 1–15.

Foster, A. and M. Rosenzweig (2010). Microeconomics of Technology Adoption. *Annual Review of Economics 2*(1), 395–424.

Gelman, A. and J. Carlin (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Pscychological Science 9*(6), 641–651.

Gilbert, B. and J. Zivin (2014). Dynamic salience with intermittent billing: Evidence from smart electricity meters. *Journal of Economic Behavior Organization 107*, 176–190.

Godlonton, S., M. A. Hernandez, and M. Murphy (2018). Anchoring Bias in Recall Data: Evidence from Central America. *American Journal of Agricultural Economics 106*, 479–501.

Greene, R., W. Timms, P. Rengasamy, M. Arshad, and R. Cresswell (2016). *Integrated Groundwater Management*, Chapter Soil and Aquifer Salinization: Toward an Integrated Approach for Salinity Management of Groundwater. Springer International Publishing Switzerland: Springer, Cham.

Gyawali, C., D. S. Dahiya, Devraj, M. Amin Bhat, and R. R. Bhandari (2016). Spatial Distribution of Physico-chemical Properties and Macronutrients in Rice Growing Soils of Haryana. *THE ECOSCAN - An International Quarterly Journal of Environmental Science 10*(3), 365–370.

Harding, M. and A. Hsiaw (2014). Goal setting and energy conservation. *Journal of Economic Behavior Organization 107*, 209–227.

Karlin, B., J. F. Zinger, and R. Ford (2015). The effects of feedback on energy conservation: A meta-analysis. *Psychological Bulletin 141*(6), 12051227.

Koundouri, P., C. Nauges, and V. Tzouvelekas (2016). Technology Adoption under Production Uncertainty: Theory and Application to Irrigation Technology. *American Journal of Agricultural Economics 88*(3).

Kumar, A., S. Rout, U. Narayanan, M. K. Mishra, R. M. Tripathi, J. Singh, S. Kumar, and H. S. Kushwaha (2011). Geochemical modelling of uranium speciation in the subsurface aquatic environment of Punjab State in India. *Journal of Geology and Mining Research 3*(May), 137–146.

Lampayan, R., R. Roderick, G. Singleton, and B. Bouman (2014). Adoption and economics of alternate wetting and drying water management for irrigated lowland rice. *Field Crops Research 170*, 95 – 108.

Lobell, D. B., A. Sibley, and J. Ivan Ortiz-Monasterio (2012, jan). Extreme heat effects on wheat senescence in India. *Nature Climate Change 2*(3), 186–189.

Lybbert, T., M. N. N, S. DJ, B. AK, and G. K (2018). Targeting technology to increase smallholder profits and conserve resources: experimental provision of laser land-leveling services to Indian. *Econ. Dev. Cult. Change 66*(2), 265306.

Macours, K. (2019). Farmers' Demand and the Traits and Diffusion of Agricultural Innovations in Developing Countries. *Annual review of resource economics 1*(1), 483.

Meenakshi, J., A. Banerji, A. Mukherji, and A. Gupta (2011). Does Marginal Cost Pricing of Electricity Affect Groundwater Pumping Behavior of Farmers? Technical report, The International Water Management Institute, Report Submitted to 3ie.

Meinzen-Dick, R., M. A. Janssen, S. Kandikuppa, R. Chaturvedi, K. Rao, and S. Theis (2018). Playing games to save water: Collective action games for groundwater management in Andhra Pradesh, India. *World Development 107*, 40–53.

Millet, P., M. Mallory, and K. Baylis (2013). The Food Corporation of India and the Public Distribution System: Impacts on Market Integration in Wheat, Rice, and Pearl Millet. *Journal of Agribusiness* (January).

Mittal, S. and V. K. Hariharan (2018). Mobile-based climate services impact on farmers risk management ability in India. *Climate Risk Management 22*, 42–51.

Okorogbona, A., L. Managa, P. Adebola, H. Ngobeni, and T. Khosa (2015). *Sustainable Agriculture Reviews*, Volume 17, Chapter Salinity and Crop Productivity. Springer International Publishing Switzerland: Springer, Cham.

Rodell, M., I. Velicogna, and J. Famiglietti (2009). Satellite-based estimates of groundwater depletion in India. *Nature 460*(20).

Russo, T. A., N. Devineni, and U. Lall (2015). *Sustainability of Integrated Water Resources Management*, Chapter Assessment of Agricultural Water Management in Punjab, India, Using Bayesian Methods. Springer International Publishing Switzerland: Springer, Cham.

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

Sekhri, S. (2014). Wells, Water, and Welfare: The Impact of Access to Groundwater on Rural Poverty and Conflict. *American Economic Journal: Applied Economics 6*(3), 76–102.

Shah, T. N., M. Giordano, and A. Mukherji (2012). Political economy of the energy-groundwater nexus in India: Exploring issues and assessing policy options. *Hydro-geology Journal 20*(5), 680–702.

Sudman, S. and N. Bradburn (1974). *Response effects in surveys: A review and synthesis.* Chicago: Aldine Publishing.

Taraz, V. (2018). Can farmers adapt to higher temperatures? Evidence from India. *World Development 112*, 205–219.

Taraz, V. and S. S. Sayre. Groundwater depletion in India: Social losses from costly well deepening. *Journal of Environmental Economics and Management 93*, 85–100.

Torres, M. M. J. and F. Carlsson (2016). Social Norms and Information Diffusion in Water-Savings Programs - Evidence from a Randomized Field Experiment in Colombia. *Economic Development Discussion Paper Series* (EFD DP 16-05).

Tsur, Y., M. Sternberg, and E. Hochman (2016). Dynamic Modelling of Innovation Process Adoption with Risk Aversion and Learning. *Oxford Economic Papers 42*(2), 336–355.

Vasilaky, K. and M. Brock (2020). Power(ful) guidelines for experimental economists. *Journal of the Experimental Science Association.*

Wok, A., L. Steg, C. Vlek, and T. Rothengatter (2015). The effects of feedback on energy conservation: A meta-analysis. *Journal of Environmental Psychology 25*(3), 273–291.

Yang, J., Q. Zhou, and J. Zhang (2017). Moderate wetting and drying increases rice yield and reduces water use, grain arsenic level, and methane emission. *The Crop Journal 5*(2), 151–158.

Figure 1: Timeline of Study Design

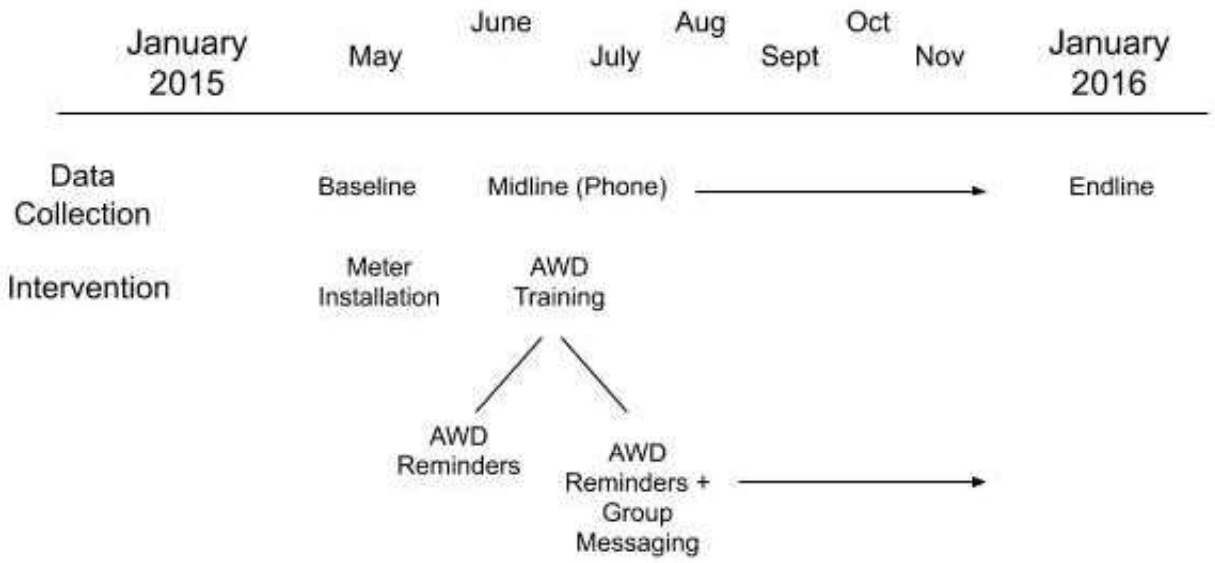

Figure 5: Weekly Total Rainfall, Average Across Villages (CHIRPS)

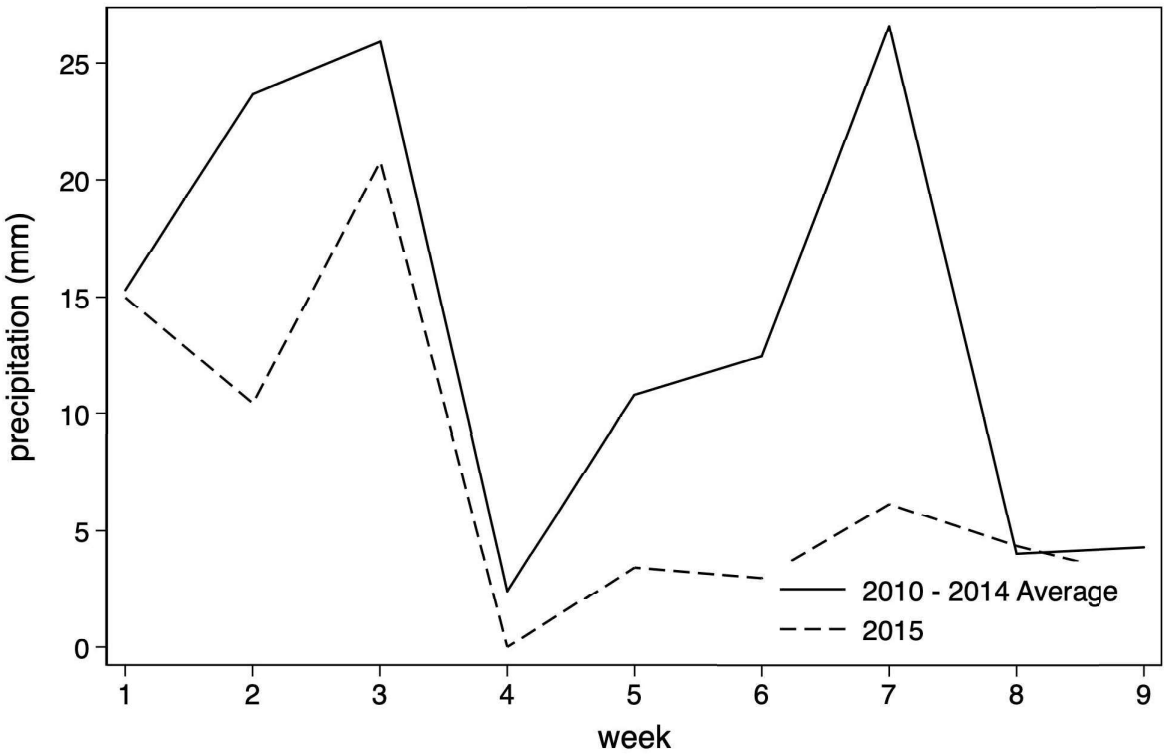

Figure 6: Weekly metered versus self-reported pumping hours per acre per week
(self-reported hours include farmers who were and were not metered)

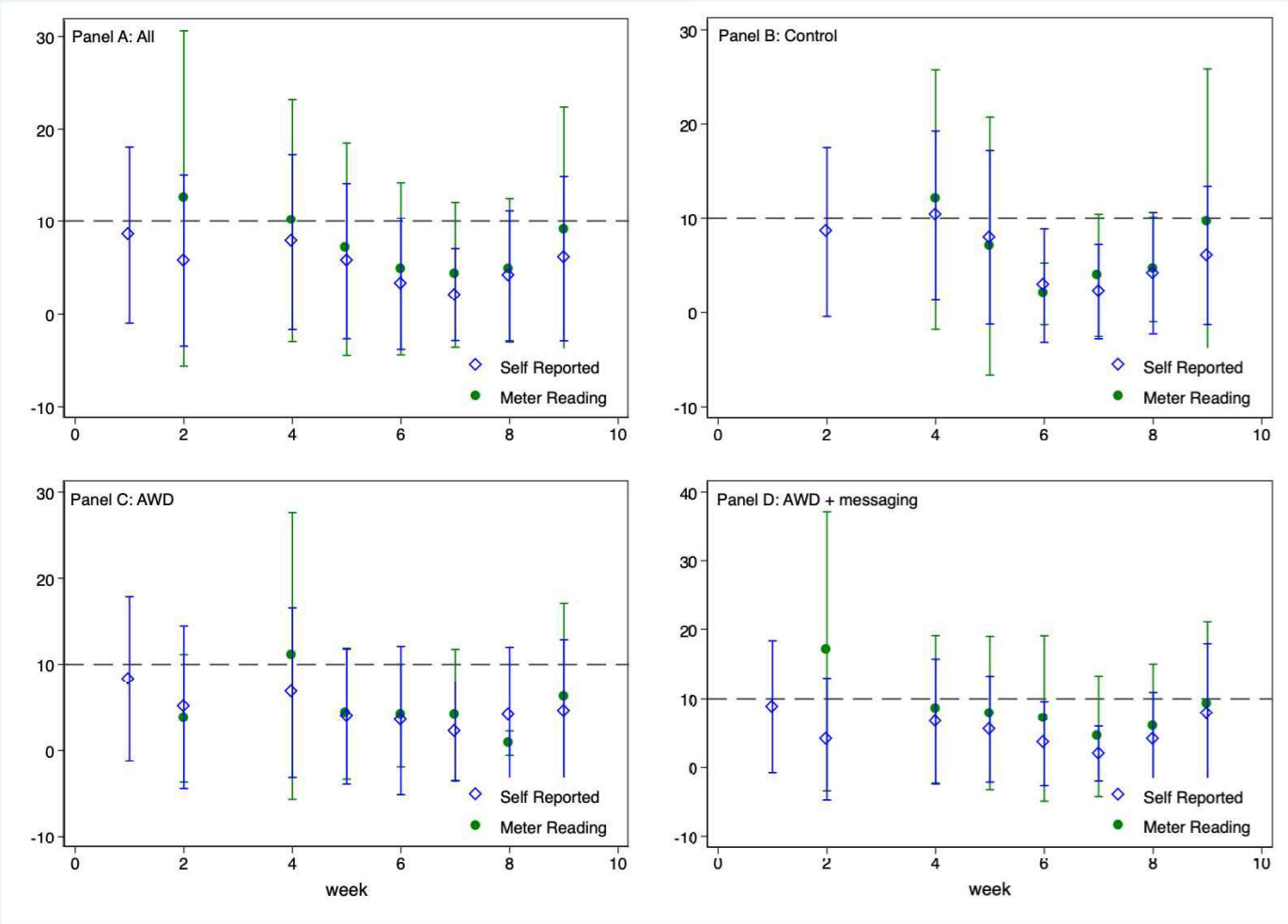

Figure 7: Weekly metered versus self-reported pumping hours per acre per week (self-reported hours include only farmers who were metered)

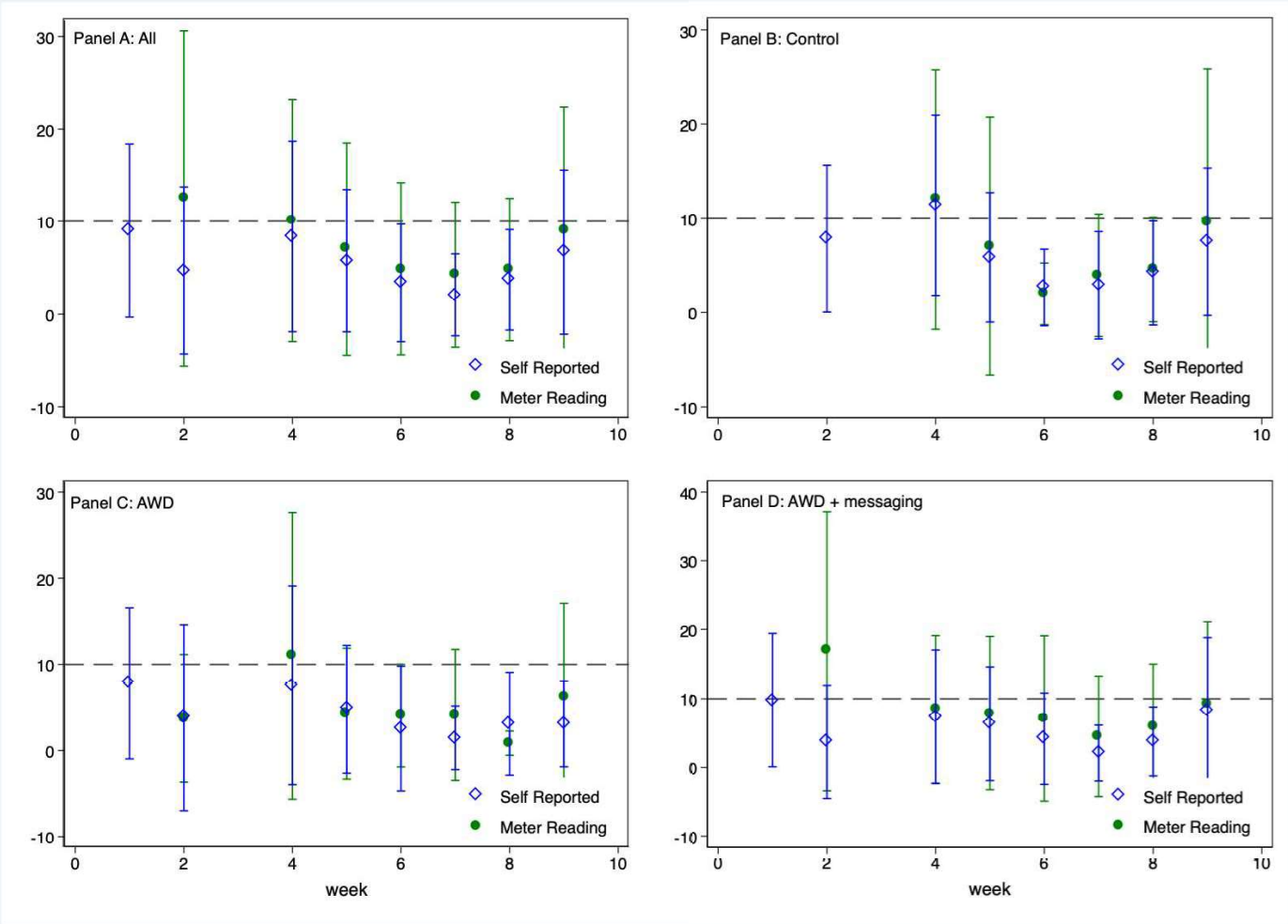

Figure 8: Weekly self-reported pumping hours per acre per week
(comparing farmers who were and were not metered)

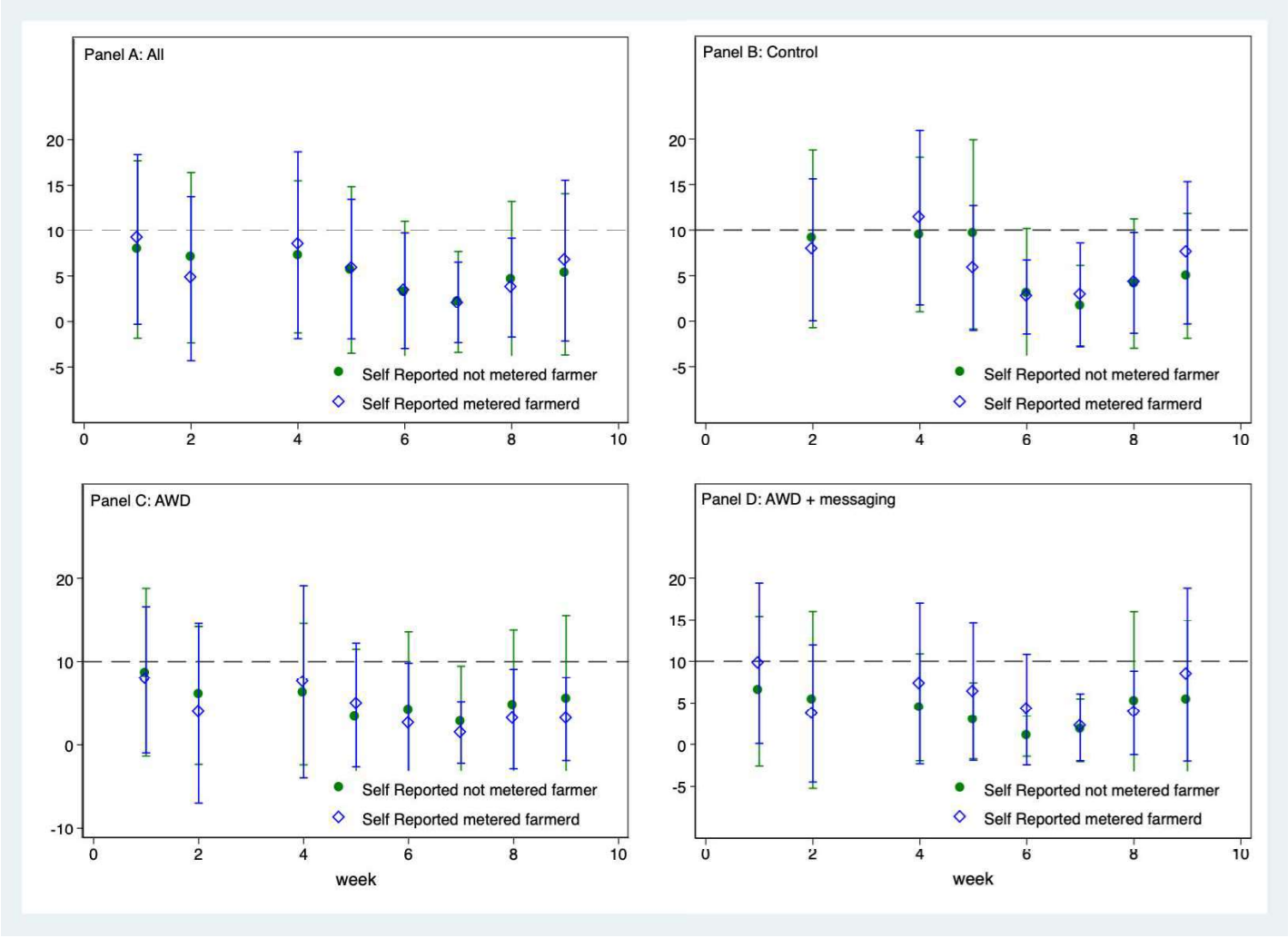

Table 1: Experimental Design Groups in Study

	Self-reported values	Values reported from metered pump
Has meter on pump	I	II
Does not have meter on pump	III	NA

Table 2: Outcomes per Farm

	Control	AWD	AWD · messaging	Prob>F
Age Respondent	46.21	42.57	47.02	0.20
	(1.43)	(1.13)	(1.28)	
Ed Respondent	8.32	9.89	8.64	0.14
	(0.47)	(0.43)	(0.47)	
If Married (y/n)	0.96	0.91	0.92	0.30
	(0.02)	(0.03)	(0.03)	
Total Land (acres)	10.49	14.66	11.72	0.34
	(0.79)	(1.96)	(1.36)	
Main Rice Plot (acres)	5.88	6.51	5.27	0.34
	(0.52)	(0.58)	(0.40)	
Year Pump Installed	2003	2004	2003	0.43
	(0.66)	(0.72)	(0.72)	
Horse Power Pump	17.36	19.72	19.49	0.16
	(0.45)	(0.58)	(0.54)	
No. Feet lowered in	56.61	51.98	59.1	0.65
last 10 years	(4.59)	(4.66)	(3.47)	
Irrigation	1.76	1.84	1.85	0.71
	(0.08)	(0.08)	(0.07)	
Water Level (ft)	137.66	155.55	142.54	0.57
	(7.32)	(9.12)	(6.78)	
Laser Level (y/n)	0.41	0.35	0.29	0.30
last 3 years	(0.05)	(0.05)	(0.05)	
Own Phone (y/n)	0.83	0.98	0.96	0.001
	(0.04)	(0.01)	(0.02)	
Own TV (y/n)	0.87	0.85	0.86	0.66
	(0.03)	(0.04)	(0.04)	
Own Cart (y/n)	0.36	0.44	0.35	0.53
	(0.05)	(0.05)	(0.05)	
Own Tractor (y/n)	0.5	0.52	0.52	0.97
	(0.05)	(0.05)	(0.05)	
Private Water Source	0.3	0.34	0.24	0.24
	(0.05)	(0.05)	(0.04)	
Observations	100	100	100	
Metered Farmers	39	42	73	
Unmetered Farmers	61	58	27	

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. Age is the age of the respondent; Education is the highest grade that the respondent completed from 1 (no education) to 13 (12th grade) or bachelor degree (14) or higher than bachelor (15); Married is 1 if married; Total land are the total land holdings for the farmer in acres; Main rice plot size is the size of the rice plot for which we surveyed; Year pump installed is the year in which the pump for the main rice plot was installed; HP pump is the horsepower of the latter pump; Lowered pump in last 10 years (f) indicates the number of feet that the pump was lowered in the last 10 years; Water Level is the number of feet below ground level the well/pump water is at; Irrigation quality is 1 if irrigation water quality has worsened, 2 if stayed the same, and 3 if improved; Laser level indicates whether the farmer laser leveled their plot in the last 3 years - an optimal time frame Aryal et al. (2015); Own tv, cart, tractor, or phone are equal to 1 if owned; Medical Expenses are total medical expenses spent in the household over the last 30 days. Private Water Source is binary value if the family has access to their own (non communal) water source.

Table 3: Baseline Summary Statistics, by Treatment Assignment and Metering
Means and Standard Errors

	Control	AWD	AWD · messaging	Metered	AWD · metered	AWD · messaging metered	Prob>F
Age Respondent	50.21	40.3	44.68	41.82	46.65	47.74	0.054
	(2.24)	(1.51)	(2.38)	(2.86)	(3.26)	(2.60)	
Ed Respondent	7.79	10.12	9.35	8.73	7.82	8.89	0.15
	(0.80)	(0.63)	(0.86)	(0.98)	(1.33)	(0.88)	
If Married (y/n)	0.97	0.91	0.87	0.95	0.94	0.96	0.21
	(0.03)	(0.04)	(0.06)	(0.05)	(0.06)	(0.04)	
Total Land (acres)	11.45	9.45	9.9	9.5	11.71	13	0.74
	(1.57)	(1.27)	(1.65)	(1.20)	(1.71)	(3.36)	
Main Rice Plot (acres)	5.31	5.81	5	6	9.18	6.04	0.77
	(0.79)	(0.87)	(0.58)	(0.76)	(1.54)	(1.10)	
Year Pump Installed	2004	2004	2003	2001	2008	2005	0.79
	(1.09)	(0.96)	(1.07)	(1.84)	(1.31)	(1.11)	
Horse Power Pump	17.18	18.58	19.4	18.5	23.53	20.44	0.43
	(0.65)	(0.83)	(1.17)	(1.27)	(1.41)	(0.85)	
No. Feet lowered in	64.59	47.43	56.7	54	59.71	66.67	0.08
last 10 years	(8.95)	(6.77)	(4.44)	(9.13)	(16.38)	(9.22)	
Water Level (ft)	137.56	145.93	130.33	151.18	194.71	152.41	0.53
	(9.44)	(9.82)	(6.40)	(19.23)	(36.74)	(16.22)	
Irrigation	1.82	2.07	1.77	1.41	1.53	1.93	0.07
	(0.12)	(0.12)	(0.11)	(0.14)	(0.19)	(0.13)	
Laser Level (y/n)	0.38	0.43	0.3	0.46	0.24	0.29	0.49
last 3 years	(0.06)	(0.07)	(0.09)	(0.08)	(0.07)	(0.05)	
Own Phone (y/n)	0.82	1	1	0.86	0.94	0.96	0.003
	(0.06)	(0.00)	(0.00)	(0.07)	(0.06)	(0.04)	
Own TV (y/n)	0.85	0.77	0.87	0.91	0.94	0.78	0.09
	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.08)	
Own Cart (y/n)	0.41	0.39	0.21	0.35	0.59	0.38	0.37
	(0.08)	(0.08)	(0.08)	(0.11)	(0.12)	(0.10)	
Own Tractor (y/n)	0.46	0.45	0.41	0.57	0.65	0.52	0.96
	(0.08)	(0.08)	(0.09)	(0.11)	(0.12)	(0.11)	
Private Water Source	0.25	0.38	0.15	0.38	0.29	0.27	0.009
	(0.06)	(0.06)	(0.07)	(0.08)	(0.07)	(0.05)	
Observations	66	70	48	34	30	52	

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. Age is the age of the respondent; Education is the highest grade that the respondent completed from 1 (no education) to 13 (12th grade) or bachelor degree (14) or higher than bachelor (15); Married is 1 if married; Total land are the total land holdings for the farmer in acres; Main rice plot size is the size of the rice plot for which we surveyed; Year pump installed is the year in which the pump for the main rice plot was installed; HP pump is the horsepower of the latter pump; Lowered pump in last 10 years (f) indicates the number of feet that the pump was lowered in the last 10 years; Water Level is the number of feet below ground level the well/pump water is at; Irrigation quality is 1 if irrigation water quality has worsened, 2 if stayed the same, and 3 if improved; Laser level indicates whether the farmer laser leveled their plot in the last 3 years - an optimal time frame Aryal et al. (2015); Own tv, cart, tractor, or phone are equal to 1 if owned; Medical Expenses are total medical expenses spent in the household over the last 30 days. Private Water Source is binary value if the family has access to their own (non communal) water source.

Table 4: Ordinary Least squares Estimates of Intention to Treat Effects on Pumping Hours (per acre per week)

	Hours from Meter		Self-Reported Hours					
			All		*Unmetered Farmers*		*Metered Farmers*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1:AWD	-3.779*	-5.129***	-1.864**	-1.526*	-1.653	-1.441	-2.407***	-2.117***
	(2.257)	(1.614)	(0.831)	(0.830)	(1.147)	(1.171)	(0.858)	(0.803)
T2:AWD · messaging	-0.348	0.141	-1.403**	-2.286***	-3.084**	-4.244*	-1.120	-1.838***
	(1.736)	(1.830)	(0.616)	(0.643)	(1.201)	(2.297)	(0.815)	(0.677)
L.above		1.060		-0.222		0.355		-0.406
		(0.968)		(0.590)		(0.885)		(0.702)
L.above · T2		-0.147		1.782***		1.410		1.873**
		(1.444)		(0.667)		(1.621)		(0.855)
Constant	25.36***	24.76***	9.851***	13.56***	9.744***	11.93***	11.03***	12.16***
	(5.824)	(6.072)	(2.424)	(2.649)	(3.769)	(2.911)	(3.375)	(3.009)
Observations	565	484	2,034	1,586	929	726	1,105	860
Number of farmers	116	116	280	271	131	124	149	147
Mean Dependent Variable	7.163	7.163	5.292	5.292	5.238	5.238	5.338	5.338

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. AWD stands for alternate wetting and drying. AWD · messaging refers to training plus text message reminders. L.above is a lagged dummy for whether the farmers' water use was above or below the previous week's village average. Columns 5 and 6 represent farmers' self-reported pumping hours, for farmers whose pumps were not metered. Columns 7 and 8 represent farmers' self-reported pumping hours, for farmers whose pumps were metered. Controls include age of the respondent, marital status of the respondent, horse power of the pump, water source (public/private), weekly total precipitation by village (source: CHIRPS). All estimates include week fixed effects.

Table 5: Input Summary Statistics, by Treatment Assignment
Means and Standard errors

VARIABLE	Control	AWD	AWD · messaging	Prob>F
Seed Expenses (Kgs/acre)	(1)	(2)	(3)	(4)
Basmati Seed	5.77	6.37	6.3	0.77
	(0.52)	(0.65)	(0.78)	
Traditional Seed	5.44	6	8.36	0.21
	(0.31)	(1.19)	(1.61)	
Labor (Days/acre)	(1)	(2)	(3)	(4)
Nursery Family	6.81	12.78	8.75	0.27
	(1.35)	(3.63)	(1.54)	
Nursery Hired	59.38	50.75	874.67	0.59
	(37.02)	(24.51)	(829.73)	
Transplantation Family	34.19	30.55	20.38	0.24
	(8.38)	(8.08)	(6.67)	
Transplantation Hired	30.61	185.25	70.33	0.19
	(7.78)	(147.46)	(18.55)	
Weeding Family	13.56	18.01	19.47	0.18
	(1.64)	(4.91)	(2.95)	
Weeding Hired	24.03	89.26	54.63	0.09
	(8.68)	(42.88)	(18.35)	
Pesticide App Family	7.35	7.02	7.44	0.98
	(1.79)	(1.52)	(1.52)	
Pesticide App Hired	35	65.08	23.07	0.46
	(17.66)	(41.93)	(10.35)	
Machinery Use (%)	(1)	(2)	(3)	(4)
Land Prep	0.47	0.45	0.45	0.95
	(0.04)	(0.04)	(0.04)	
Transplantation	0.16	0.09	0.09	0.26
	(0.04)	(0.03)	(0.03)	
Weeding	0.02	0.04	0.04	0.56
	(0.01)	(0.02)	(0.01)	
Pesticide Application	0.07	0.08	0.08	0.90
	(0.02)	(0.02)	(0.02)	
Fertilizer and Pesticide (Kgs/acre)	(1)	(2)	(3)	(4)
Potash	30.01	25.82	15.65	0.20
	(8.15)	(7.00)	(3.01)	
Urea	85.93	82.98	92.97	0.81
	(5.33)	(8.48)	(14.01)	
Pesticide Use (%)	0.9	0.89	0.89	0.95
	(0.02)	(0.02)	(0.02)	

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. Family refers to the number of days family members supplied to the tasks of preparing the nursery, transplanting rice seedlings, weeding, and applying pesticides. Family refers to the total number of days hired for the tasks of preparing the nursery, transplanting rice seedlings, weeding, and applying pesticides. Machinery use (%) refers to the percentage of respondents who use machines for land preparation, transplantation, weeding, and pesticide application.

Table 6: Ordinary Least Squares Estimates of Treatment Effects on Yields and Paddy Consumption

	Yield (kgs/acre)	Consumption last 30 days (kgs)
Panel A: All Farmers	(2)	(4)
AWD	1,107	0.604
	(2,307)	(1.520)
AWD · messaging	-1,351	3.220*
	(1,413)	(1.758)
Constant	5,552	6.519
	(6,887)	(6.437)
Observations	238	240
Mean Dependent Variable		
Panel B: Unmetered Farmers	(2)	(4)
AWD	3,650	2.741
	(4,171)	(1.904)
AWD · messaging	-134.3	3.005*
	(1,387)	(1.506)
Constant	-4,482	2.823
	(5,691)	(5.375)
Observations	116	118
Mean Dependent Variable		
Panel C: Metered Farmers	(2)	(4)
AWD	-3,202	-3.023
	(2,526)	(2.740)
AWD · messaging	-4,690	0.696
	(3,280)	(2.892)
Constant	-60,804	107.1
	(391,001)	(204.5)
Observations	122	122
Mean Dependent Variable	5029	11.16

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. AWD stands for alternate wetting and drying. AWD · messaging refers to training plus text message reminders. Metered refers to farmers randomly selected to have their pumps metered for pumping use. Yield refers to total rice harvest in kilograms divided by rice acreage for the main rice plot. Consumption refers to total household kilograms of rice consumed in the last 30 days. Controls include age of the respondent, marital status of the respondent, and whether the respondent owned their own phone. Since this is a cross-sectional regression using endline data, there are no week fixed effects nor weekly precipitation controls.

Table 7: Probit Estimates of Marginal Treatment Effects on Perceived Cooperation and Trust

	Cooperate	Trust
Panel A: All Farmers	(2)	(4)
AWD	0.326**	0.0694
	(0.155)	(0.271)
AWD · messaging	0.449**	-0.275
	(0.197)	(0.283)
Constant	0.0182	0.111
	(0.674)	(0.655)
Observations	249	249
Mean Dependent Variable		
Panel B: Unmetered Farmers	(2)	(4)
AWD	0.161	0.356
	(0.233)	(0.496)
AWD · messaging	0.601	-0.0985
	(0.461)	(0.548)
Constant	1.456***	-0.487
	(0.439)	(0.876)
Observations	121	117
Mean Dependent Variable		
Panel C: Metered Farmers	(2)	(4)
AWD	0.766*	-0.240
	(0.441)	(0.392)
AWD · messaging	0.272	-0.424
	(0.300)	(0.379)
Constant	1.111***	0.106
	(0.193)	(0.558)
Observations	123	123
Mean Dependent Variable	0.922	0.906

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. AWD stands for alternate wetting and drying. AWD · messaging refers to training plus text message reminders. Metered refers to farmers randomly selected to have their pumps metered for pumping use. Cooperation refers to t/f responses to this statement "The farmers in my community are cooperative with one another." Trust refers to t/f responses to this statement "I trust the other farmers in my community to do what is best for the whole community." Controls include age of the respondent, marital status of the respondent, and whether the respondent owned their own phone. Since this is a cross-sectional regression using endline data, there are no week fixed effects nor weekly precipitation controls.

Table 8: Ordinary Least squares Estimates of Intention to Treat Effects on Pumping Hours (per acre per week)
Wild Bootstrap Standard Errors

	Hours from Meter		Self-Reported Hours					
			All		*Unmetered Farmers*		*Metered Farmers*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1:AWD	-3.004	-3.173*	-1.793*	-1.550***	-1.622	-1.318	-2.346**	-2.144***
	(0.195)	(0.0552)	(0.934)	(0.484)	(1.394)	(0.926)	(1.111)	(0.696)
T2:AWD·messaging	0.700	0.988	-1.357**	-1.894***	-2.870**	-3.884*	-1.179	-1.513*
	(0.530)	(0.518)	(0.692)	(0.655)	(1.429)	(2.054)	(1.046)	(0.793)
L.above		2.756*		-0.0739		-0.184		0.365
		(0.0712)		(0.517)		(0.815)		(0.708)
T2·L.above		0.870		1.053		1.570		1.010
		(0.669)		(0.839)		(2.262)		(0.964)
Constant	23.93***	17.00***	10.35***	12.62***	5.033	6.675*	12.59***	10.24***
	(0)	(0.00144)	(0)	(1.954)	(3.751)	(4.012)	(0)	(2.265)
Observations	565	484	2,034	1,586	929	726	1,105	860
Mean Dependent Variable	7.163	7.163	7.163	7.163	5.238	5.238	5.338	5.338

*** p<0.01, ** p<0.05, * p<0.1 Standard errors, clustered at the village level, reported in parentheses. AWD stands for alternate wetting and drying. AWD · messaging refers to training plus text message reminders. L.above is a lagged dummy for whether the farmers' water use was above or below the previous week's village average. Columns 5 and 6 represent farmers' self-reported pumping hours, for farmers whose pumps were not metered. Columns 7 and 8 represent farmers' self-reported pumping hours, for farmers whose pumps were metered. Controls include age of the respondent, marital status of the respondent, horse power of the pump, water source (public/private), weekly total precipitation by village (source: CHIRPS). All estimates include week fixed effects.