Measuring demand for groundwater irrigation: Experimental evidence from conservation payments

Nick Hagerty and Ariel Zucker

April 17, 2018

Note to the Reader

This paper is intended for submission to a journal as a "registered report," for which peer review takes place prior to data collection and analysis. Many journals in other disciplines already have adopted this publication process; in economics the *Journal of Development Economics* recently announced a pre-results review process. Initial publication decisions are made on the basis of a prospective plan for an empirical project, which should include the standard elements of a pre-analysis plan. If accepted, we would submit a full paper following the completion of the analysis, which would be judged only according to whether the data collection and analysis were executed as specified. This pilot process is aimed at increasing transparency in research and reducing publication bias, and we are excited to be early participants in this experiment. Further information on registered reports at the *Journal of Development Economics* is available from the World Bank blog post announcing the program and the Guidelines for Authors.

1 Introduction

Groundwater is a major source of irrigation and drinking water worldwide, especially for farmers in developing countries (Ministry of Agriculture of the Government of India, 2014). Unfortunately, falling groundwater levels are creating negative consequences in many regions. Depletion reduces water availability, raises the cost of further extraction, may harm water quality, and can increase poverty and conflict (Sekhri, 2014). While groundwater pumping is currently unregulated in much of the world, many regulatory tools are available, ranging from quantity restrictions and tradeable quotas to simple price instruments. However, to implement these tools efficiently, a regulator requires knowledge of the demand for groundwater – a key input for which evidence is thin.

This paper presents an experimental protocol to measure the price response of demand for groundwater in irrigated agriculture. To measure demand, we introduce a mechanism that implements a price incentive without requiring the power of taxation: payments for reduced groundwater pumping, or "conservation credits". This study will be conducted as a randomized controlled trial among well-owning farmers in the Indian state of Gujarat between 2018 and 2020. The basic research design will be to (1) install meters on the groundwater pumps of all study participants, (2) offer payments for reduced pumping, relative to a benchmark quantity, to a randomly selected sub-sample of participants, and then (3) compare the quantity of groundwater extracted by these farmers to that of the rest of the sample (i.e., the control group).

Besides allowing us to measure demand, our intervention may be a promising policy tool in itself. By offering payments for voluntary conservation, we may be able to overcome constraints often faced in regulating common-pool resources in developing countries. One such constraint is weak enforcement capacity: In many areas, a natural regulator for groundwater would be the state-owned utility that provides the electricity used for pumping – but consumer-level metering is rare, and electricity theft is widespread (Antmann, 2009; Northeast Group LLC, 2014; Golden and Min, 2012). Another constraint is political concerns: Both energy subsidies and open access to groundwater are often entrenched means of redistribution; in India, reform efforts are commonly met with forceful protests (Sovacool, 2017).

The conservation credits model may be able to relax these constraints in two ways. First, our program may be easier to enforce than electricity sales; both technical and institutional features of the program may make cheating both more difficult to do and easier to detect. Second, we do not attempt to interfere with existing *de facto* entitlements. Unlike (for example) a new Pigouvian tax on groundwater consumption, which has large costs to large users of free groundwater, we instead offer payments relative to existing usage patterns. While conservation credits require large expenditures, a Pareto improvement may be possible: an electric utility may be willing to implement conservation credits if the outlay per unit of energy conserved is smaller than the marginal cost of electricity provision.

Our analysis will consist of three parts. First, evaluating our intervention as a whole, we will measure how much water and energy is saved by the conservation credits program. This will provide reduced-form evidence on the response of demand for groundwater irrigation to price incentives, as well as evidence on the ability of the conservation credits model to reduce resource consumption in our context. Our primary outcome is duration of pump operation, as measured directly using hours-of-use meters. We will also estimate treatment effects in energy and water equivalents, and assess mechanisms of water conservation and follow-on environmental and economic impacts. To assess "leakage" in this program (negative spillovers to non-monitored actions), we will also estimate intervention effects on the use of other, unmetered water sources.

Second, we will use the design of our intervention to estimate the slope of groundwater demand with respect to price, a parameter that is an important input to the design of any type of groundwater regulation. We will estimate demand by instrumental variables, using treatment group assignment as instruments for price faced at the margin. To obtain a quantitative estimate of demand parameters, we need an instrumental variables approach; intent-to-treat estimates do not suffice because of the structure of our intervention design. Specifically, it would be cost-prohibitive to offer enough incentives to ensure everyone is marginal – in practice, some participants find their benchmark too low to affect their decisions. Additional random variation in both prices and benchmarks will help to increase statistical power and first-stage instrument strength.

Third, we will assess the cost-effectiveness of the intervention as implemented in the study. Conservation credits could be implemented by even a budget-constrained electric utility if the cost of the energy conserved is larger than the cost of the program. To evaluate the viability of this potential Pareto improvement, we will compare these costs and test whether an electric utility could be enlisted to reduce groundwater consumption using conservation credits. If the answer is no, we will calculate the minimum per-unit groundwater conservation subsidy that would be required for conservation credits to yield net benefits.

Our study design is informed by a small pilot trial implemented among 90 farmers in our study region during the winter of 2017-18. This pilot demonstrates logistical feasibility of our intervention: farmers approached were overwhelmingly willing to participate in the study and install meters, the meters functioned properly, and we observed little evidence of tampering. The pilot also yielded several improvements in intervention design, as well as preliminary data on pumping hours that informed our sample size calculations. Results from the pilot are highly imprecise but point estimates are consistent with a large reduction in pumping hours in the treatment group. Our study also follows an earlier, non-randomized pilot of a similar program in northern Gujarat (Fishman et al., 2016).

This study will make several contributions. First, this study will provide, to our knowledge, the first experimental evidence on the price sensitivity of demand for groundwater. Price variation is scarce for an open-access resource, so most previous estimates have used proxies for the cost of pumping (Gonzalez-Alvarez et al., 2006; Hendricks and Peterson, 2012), but these proxies may be correlated with other determinants of groundwater demand. Bruno (2018) exploits panel variation in prices across three regions of an irrigation district in California, but there is still a possibility that these prices may have responded to groundwater consumption; an experiment can rule out both concerns. We also focus on a developing country, where evidence on groundwater demand is particularly scarce. Meenakshi et al. (2013) use differences-in-differences to study a phased-in switch to metering in West Bengal, India, but they rely on self-reported pumping data and find imprecise results. Badiani and Jessoe (2017) estimate an aggregate price elasticity using panel variation in the fixed cost of an electricity connection, but marginal incentives may produce quite different results.

Second, we will contribute to the literature on the cost-effectiveness of payments for environmental services (PES) for resource conservation. Our conservation credits intervention has the same basic structure as hundreds of programs designed to incentivized the provision of environmental services, ranging from increased forest or wetland cover, to reduced input intensity in agriculture.[1] Despite their prevalence, rigorous evaluation of these types of programs has been limited (see Pattanayak et al. (2010) and Börner et al. (2017) for reviews). Most existing evaluations use covariate matching and are unable to address selection bias, a particular concern for a voluntary program. The two exceptions are Jayachandran et al. (2017), who use a randomized controlled trial to find that conditional payments to forest-owning households in Uganda reduce deforestation rates by 50 percent, and Jack and Cardona Santos (2017), who find that while contracts for tree planting in Malawi increase trees planted, they also increase tree clearing on unenrolled plots. Our study will provide evidence on the feasibility and effectiveness of PES in a novel context: promoting irrigation efficiency in agriculture. We will also build on previous studies by using detailed survey data to investigate the behavioral mechanisms underlying the response to a PES program. Understanding the response to this program will inform our understanding of whether a PES model holds promise as a method for governments and donors to reduce energy and water use.

Finally, we will contribute to literature connecting the price response of electricity consumption in developing countries to policy decisions about energy-sector investment and reform. Experimental and quasi-experimental studies are still limited, but a few have been conducted recently on rural households in Columbia (McRae, 2015), urban households in South Africa (Jack and Smith, 2016), and new grid connections in Kenya (Lee et al., 2018).

2 Background

2.1 Optimal groundwater policy: A framework

Groundwater is a shared, common-pool resource. Extraction by one user (most often irrigators) imposes an externality on other users in the form of lower water availability and higher costs of extraction. Multiple regulatory tools - including both quantity and price instruments - are available to reduce over-extraction and restore efficiency, and demand for groundwater is an essential input to all of them. In this section we show how the optimal Pigouvian price level is set, and how this calculation is affected by the demand for

[1] For example, in the United States alone, payments are available to farmers for actions to mitigate flood and wildfire risks, provide habitat for endangered species, salinity mitigation, and water and energy conservation.

groundwater. We focus on price regulation because our study implements a type of price instrument, but the analysis would be similar for the quantity instruments more frequently used for groundwater management.

Figure 1 illustrates consequences of price regulation in the presence of groundwater externalities. Irrigators have aggregate inverse demand for groundwater as a function of water quantity, $D(q)$. Inverse demand equals private marginal benefits net of private marginal costs of extraction; it first declines with quantity but eventually slopes upward as marginal costs rise. Extraction generates social marginal damages, $SMD(q)$, which increases with quantity. Although this analysis represents the situation at a single point in time, it can fully incorporate dynamics: the present discounted value of future costs of today's extraction may be included in demand (the internalized portion) and social marginal damages (the remainder).

When groundwater extraction has a price of zero, irrigators continue using water until net private marginal benefits are zero - where the demand curve intersects the x-axis, or q_0. This level of extraction is inefficient, since the social marginal damages are greater than the net private marginal benefits. The efficient level of extraction, instead, is found where these two curves intersect, or q^*.

One way to achieve this allocation is through a price, or tax, per unit quantity extracted. If the price p is set to equal p^*, the value of social marginal damages at q^*, irrigators will fully internalize the externality of extraction, shifting down the effective demand curve. Then, they will extract only up to the efficient quantity q^*, since net private marginal benefits including the tax are zero. To set this per-unit price p^*, a common heuristic is to set the price equal to the social marginal damages as measured locally. If social marginal damages are constant, the slope of demand does not matter, since the efficient quantity is simply whatever amount results from this price.

However, there are two reasons a policymaker pursuing price regulation may need to know the full shape of the groundwater demand curve. First, social marginal damages may not be constant. In Figure 1, if the price were set at $SMD(q_0)$, the resulting quantity extracted would be far too low. Constant social marginal damages may be a reasonable approximation over the range of groundwater conserved in small programs in large aquifers, but the slope of the demand curve is essential for larger programs or smaller aquifers. Second, even if social marginal damages are constant, the process of enacting a new policy may incur costs (such as political or administrative costs). Whether the policy is worthwhile depends on the quantity of water conserved, which can only be predicted with knowledge of the demand curve.

2.2 Existing evidence: Costs, benefits, and damages of groundwater extraction in irrigated agriculture

Existing evidence is relatively thin for both the social damages and demand functions for groundwater extraction. Social damages are difficult to quantify overall, but the components are well understood. Some components have known values, while others are best estimated using scientific models. Demand for groundwater, which is the difference between private marginal benefits and private marginal costs, is less well understood. Private marginal costs can be modeled fairly easily, but private marginal benefits are unknown. Our study fills the gap in knowledge by directly estimating demand.

2.2.1 Social damages

Social damages from groundwater extraction come first through the depletion of the resource. Groundwater extraction by one user generally leads directly to a decline in water levels for other users. The precise relationship between extraction and water levels depends on geology, topography, soil, rainfall, and climate. Deeper groundwater levels raise the cost of extraction, which can lead to increases in poverty and conflict (Sekhri (2014)). Depletion can also degrade water quality, either through inherent local properties of soil and geology, or by drawing in seawater from the ocean in coastal areas.

These externalities can be complex and difficult to estimate, since the spatial extent of the extraction externality varies greatly across locations. Depending on geology, in some areas, the externality may fall almost entirely on a small group of neighbors, in which case Coasian bargaining may sometimes be able to govern the aquifer efficiently. However, in many areas, and especially over longer periods of time, the externality is felt over a very large area, making local cooperation less likely to be sustained.

Another major source of social damages, which is easier to measure, is the costs associated with the energy required to pump groundwater to the surface. Typical energy sources are electricity and diesel, both of which create greenhouse gases and air pollution. In many developing countries, including almost all states of India, political pressure constrains governments to provide electricity to agricultural customers at a marginal price of zero. In this case, the social marginal damages of groundwater extraction include the marginal cost of electricity provision by the electric utility.

2.2.2 Demand

Private marginal costs in the short run can be modeled reasonably easily: they depend on the price of fuel (which may be zero), water levels, and pump characteristics. In the long run - that is, over large changes in water levels - discontinuities in private marginal costs may arise from deepening wells or purchasing new pump hardware.

Private marginal benefits of groundwater extraction are more difficult to estimate since they depend on the agricultural production function and any non-profit-maximizing behavior by farmers. Anecdotal evidence suggests that, especially in developing countries, water inputs often exceed yield- and profit-maximizing levels. Instead of measuring inputs precisely, some farmers simply flood their fields - which would suggest that private marginal benefits are low at the current equilibrium. Because these private benefits are difficult to model, we instead directly estimate groundwater demand using a revealed-preference approach, in which we observe how quantity extracted changes with price.

2.3 Conservation credits as a Pigouvian tax

Our objective is to estimate groundwater demand by varying the price of extraction. As an external party lacking the power of the state, we cannot require irrigators to pay a tax. Instead, we offer payments for reduced water extraction, relative to a benchmark amount - an intervention called "conservation credits." This intervention provides the same marginal incentives as a Pigouvian tax, at least for some participants.

Figure 2 illustrates the budget set of the conservation credits contract. Two thresholds are set: a benchmark, and a maximum payment. If the irrigator extracts a greater quantity than the benchmark, the payment is zero. If the irrigator conserves water relative to the benchmark, the payment equals the price times the difference between the quantity and the benchmark. If the irrigator conserves very large amounts of water, the maximum payment may be reached, after which further conservation does not increase the payment.

Under a Pigouvian tax, all irrigators are marginal to the incentive, in the sense that any positive quantity extracted is subject to a per-unit price. Under conservation credits, many irrigators are marginal, but not all. To see this, Figure 2 plots quasi-linear indifference curves over groundwater extraction (including both the private benefits and costs) and payments of conservation credits. Without conservation credits, the budget set is flat and coincides with the x-axis; with conservation credits, the budget set is piecewise linear. Irrigator A is marginal: her indifference curves are tangent to the x-axis at q_0^A and tangent to the conservation credits budget set at q_1^A, indicating that she will reduce groundwater extraction when eligible for conservation credits. Irrigator B is extra-marginal: his indifference curves are tangent to both budget sets at q_B, indicating that he will not reduce extraction in response to conservation credits.

2.4 Equivalence between water, energy, and pumping time

Some of the costs and benefits discussed so far are in units of water quantity, while others are in units of energy consumed. In data collection, our main outcome of interest will be a third unit: pumping time. These three objects are closely related and can be converted using mechanical formulas:

$$E = \frac{P_b}{\eta_m} \times t \qquad q = \frac{P_b \eta_p}{kh} \times t \tag{1}$$

where E is energy consumed, q is quantity of water pumped, and t is duration of pump operation. P_b is the power rating of the pump's motor ("brake horsepower"). η_m and η_p are the motor and pump efficiencies; they are unitless, between zero and one. h is the total hydraulic head, approximately equal to the depth to water level (plus friction and outlet pressure). k is a conversion constant, equal to 3960^{-1} when q is in gallons, P_b is in horsepower, t is in minutes, and head is in feet.

Slopes of demand can be related similarly by differentiating both sides of the formulas in Equation 1. The price elasticities of demand for water, energy, and hours are equal.

3 Study Setting and Experimental Design

To estimate groundwater demand, we will implement a randomized controlled trial among groundwater-irrigating farmers in Gujarat, India. The trial will have two overarching treatment arms: *conservation credit* farmers will be eligible to receive payments for conserving groundwater below a benchmark, whereas *control* farmers will receive no such incentives.

3.1 Setting

Our trial will be implemented in Saurashtra, a water-scarce region of Gujarat state, India. The study villages are located in coastal areas of Talaja block in Bhavnagar district, where falling groundwater levels lead not only to increased irrigation costs, but also to increased risk of seawater intrusion into the freshwater aquifer. Salinity levels in Talaja aquifers are already extremely high (Central Ground Water Board (2013)), with 60% of villages either prone to increased salinity or already partially or fully saline,[2] reducing the ability of farmers to grow high-value crops (Samadhan E Cube Innovator Pvt. Ltd. (2016)).

The primary source of employment in Talaja block is in agriculture (Registrar General and Census Commissioner of India (2001)). The literacy rate is low compared to the rest of India, at approximately 44%. In

[2] Prone-to-saline indicates average total dissolved salt (TDS) concentration >500mg/L, partially saline from 1000 to 2000 mg/L, and fully saline > 2000 mg/L.

addition, households are relatively large, with an average household size of 5.7 individuals (compared to 4.7 across India). Only 3% of individuals are from Scheduled Castes or Scheduled Tribes. Agricultural land is primarily irrigated by groundwater (47%), although 28% is surface-water irrigated, and the remaining 25% is rain-fed (Registrar General and Census Commissioner of India (2011)).

3.2 Enrollment

We will recruit our sample from 44 study villages. The study villages were selected by our implementing partner, the Coastal Salinity Prevention Cell (CSPC). CSPC is not yet working in these villages, but plans to roll out a number of programs beginning in 2018, ranging from health to agricultural development interventions.

The sample in each village will be randomly selected from eligible households of landowning villagers who are willing to participate, following the procedure developed in our pilot study. The sample frame will be formed on the basis of official village landowner lists, which can be obtained from the village *talati* (accountant). These lists include all land-owning villagers as of the date of the list (and sometimes include other information, such as landholding size and location). CSPC will augment the village list with phone numbers[3] through its network of village extension volunteers (local villagers who carry out simple organizational tasks for a small stipend).

Random sampling will be conducted as follows. First, each name on a village list will be assigned a random number. Surveyors will call and/or visit individuals on the list, in the order of the number assigned, to determine if the primary agricultural decision-maker (PAD) in the household meets the study eligibility criteria, and is willing to participate in the study. Surveyors will then visit the eligible and willing PADs to obtain informed consent for enrollment in the study.

In order to be eligible for the study, the household's PAD must meet the following criteria:

Inclusion Criteria

- Also be the primary agricultural decision maker for the land owned by the household.
- In the previous Rabi (winter) season:
  - Must have planted crops.
  - Must have irrigated at least one farm using primarily groundwater.
  - Must have solely used an electric-powered pump for this purpose.

[3] A survey funded by CSPC in Talaja and neighboring Gogha block in 2016 revealed 78% of farming households owned at least one phone.

- Must plan to farm and irrigate during the next Rabi season.

Exclusion Criteria

- May not have a diesel pump in use on the primary water source.
- May not have multiple pumps, or pump starters, in use on the primary irrigation source (i.e. well).

Farmers who meet the eligibility criteria, complete a baseline survey, and consent to the full study will be enrolled.

3.3 Sample size

We plan to draw a total sample of 2,200 well-owning farmers, equally divided between treatment and control groups. These choices are based on power calculations using meter reading data from a pilot of 90 farmers (the pilot is more fully described in Section 6). Our primary object of interest is the intent-to-treat estimate: the average treatment effect of eligibility for conservation credits on the duration of pump operation. We measure this as the difference in group means of total pumping hours after partialing out strata and month effects. This is likely conservative, since individual-level time-varying covariates may be able to improve efficiency.

First, we divide the sample between treatment groups equally, since the primary object of interest is a simple difference in means. Assuming equal variance in the outcome variable across treatment groups (which our pilot data cannot reject), power is maximized at a treatment proportion of 0.5.

Second, we calculate the sample size required for a minimum detectable effect of 10 percent of the sample mean pumping hours, at a power of 0.9. This is the sample size required to reject a null hypothesis of no effect with 90 percent probability when the true effect is 10 percent of the mean. We choose 10 percent because it is a salient and quantitatively reasonable threshold. From pilot data, the conditional variance of total pumping hours (i.e., after partialing out several baseline covariates) is 0.47. The required sample size is then 1,990. Allowing for an approximately 10 percent attrition rate,[4] the total sample to be recruited is 2,200.

Figure 3 plots full power curves using the pilot data for a range of sample sizes and minimum detectable effects.

[4]This is larger than the 2 percent attrition in the pilot; we are conservative about attrition due to the increased duration of the planned experiment compared to the pilot, to 18 months from 5 months.

3.4 Randomization

Randomization will be stratified by village and forecasted hours of irrigation. Specifically, the final sample within each village will be divided into above- and below-median forecasted hours of irrigation, creating two equally-sized cells in each village. Farmers in each cell will then be randomly allocated to one of two treatment arms using a pseudo-random number generator (Stata software): 50% will be allocated to the Conservation Credit arm, and 50% to the Control arm. Within the Conservation Credit arm, we will cross-randomize both the benchmark below which individuals are incentivized between a high and low option, and the size of the marginal conservation incentive between a high and low option, resulting in four equally-sized sub-treatments (Figure 4). The intervention will run for approximately 18 months, from randomization in late summer 2018, until the final meter readings are completed in February, 2020.

3.5 Interventions

Conservation credits

Participants in the Conservation Credits arm will have an hours-of-use meter installed on the electric pump starter of their primary irrigation source. The meter measures the total hours of irrigation done by the farmer. Meters will be read monthly by CSPC village extension volunteers.

Farmers will be incentivized for conserving water for five months of the Rabi season, from September-January, across two consecutive years. This is the period of peak irrigation; as there is typically no rainfall during Rabi, agriculture is entirely dependent on irrigation. At each meter reading, farmers are informed of their benchmark for the following month, and the payment for the previous month is calculated. Payments are awarded at a fixed rate for consuming fewer hours of irrigation than the monthly benchmark, according the formula:

$$\text{Payment}_{it} = \max\left(0, \text{price}_i \times ((\text{hours benchmark})_{it} - (\text{hours consumed})_{it})\right) \tag{2}$$

where price_i is the per-hour incentive rate, $(\text{hours benchmark})_{it}$ is an individual-month-specific benchmark, and $(\text{hours consumed})_{it}$ is the monthly meter reading. The payments are later disbursed as checks.

Conservation Credit Sub-treatments

The four Conservation Credits sub-treatments differ along two dimensions: the per-hour incentive rate, and the benchmark. Individuals assigned a *high price* receive 40 INR (0.61 USD) per hour conserved, and those assigned a *low price* receive 20 INR (0.31 USD) per hour conserved. The prices were chosen to be realistic estimates of the groundwater price that a policymaker might wish to set. The prices are similar to the cost of electricity provision for the median farmer, which is approximately 30 INR per hour of use (authors' calculation).[5]

Individuals assigned the *high* and *low benchmark* receive 125% and 75%, respectively, of their forecasted monthly hours of irrigation. The monthly hours of irrigation forecast used to set benchmarks (and to stratify randomization) will be created from baseline survey data collected before the program is introduced. The survey collects the self-reported duration of irrigation in the previous year's Rabi season by asking for the number of irrigations made during Rabi, the average duration of each irrigation, and the first and last irrigation dates. This method of constructing benchmarks is potentially liable to manipulation: if farmers know how survey answers will be used, we might expect farmers to artificially inflate their reported irrigation to increase their expected conservation credit payment. Because farmers will not yet know the program details at baseline, we do not foresee manipulation in our setting. However, any scaled-up program would have to rely on a different method of benchmark setting, such as collecting actual usage data.

Control

Participants in the Control arm will also have an hours-of-use meter installed and read monthly for 18 months. However, these farmers will not be incentivized for conservation.

4 Data

4.1 Data collection

In order to conduct our analysis, we will collect four datasets. First, we will conduct a baseline survey with both self-reported and field measurement components prior to randomizing participants into treatments. Self-reported data will include demographic and socioeconomic characteristics, such as landholding size and household size; cropping, crop management, and irrigation decisions in the previous year; the power of the primary pumpset, and water conservation strategies and attitudes. Field measurements will include the precise geolocation, depth-to-water and salinity levels (i.e., total dissolved solids) of each well on the participant's largest farm. All data will be collected electronically through tablet surveys.

Second, we will directly measure groundwater pumping for all study participants, using hours-of-use meters installed on the pump starter of each participant's primary irrigation source.[6] Village extension

[5] A rate of 30 INR per hour is approximately equal to the unsubsidized average cost of electricity supply in Gujarat for the power rating of a typical pumpset in the pilot region. That is: (5 INR/kWh average cost of electricity provision in Gujarat) * (6.2 HP average pump brake power) / (74% typical motor efficiency) * (0.75 kW/HP conversion factor) $\approx$ 31 INR/hr.

[6] Digital hours-of-use meters manufactured by International Instrument Industries (model: Selec, LT-920).

workers will record meter readings each month using a digital tablet survey. Meter data quality will be assured through random audits, in which a research associate will compare the digitally recorded meter readings with dated, geo-located photographs of the meter dial included on the tablet survey.

Our third and fourth datasets will be collected in two endline surveys, one after each Rabi season. These two tablet-based surveys will collect the same field measurements and retrospective self-reported information as the baseline survey, as well as specific questions on changes to irrigation behavior and new technology adoption.

Finally, we will collect a supplementary dataset of water and energy consumption, measured simultaneously with hours of use. We will use this dataset us to calibrate pump and motor efficiencies under realistic conditions similar to our study sample, which in turn will allow us to convert between our measured hours of irrigation and the water and energy equivalents. For each major pump and motor type in our sample, we will select small calibration sub-samples and compare hours of irrigation with readings from both a portable electricity meter and an ultrasonic water flow meter.

4.2 Data processing

To assess the response of groundwater irrigation to water prices, our primary outcome will be monthly hours of groundwater irrigation. Hours of irrigation during each meter-reading period will calculated as the difference between total hours consumed at the end and beginning of the period. For individuals whose meters have been disconnected following the drying of a well, hours will be recorded as usual (i.e. according to the meter dial). For individuals whose meters are otherwise tampered with (e.g. if the meter is disconnected or broken but the well is not dry), hours will be recorded as missing. Because meter-reading periods may vary slightly over time and across individuals, we will normalize the measured hours of irrigation in each period by the number of days in the period.

We will construct two secondary outcomes, water consumption and energy consumption, from hours of irrigation using the conversion formulas in Equation 1. Pump and motor efficiencies will be imputed based on pump and motor type from the dataset of pump and motor efficiencies, pump power will be collected in surveys, and monthly depth-to-water will be interpolated from baseline and endline measurements.

Other secondary outcomes will be measured in each of the two endline surveys. One group of outcomes measures the environmental and economic follow-on effects of water conservation. We will assess environmental impacts through measured water depth and salinity levels in the metered wells, and economic impacts through self-reported crop yields for selected crops, crop revenue, and farm profits in the Rabi season. Crop yields will be measured through questions asking the total kg of crops harvested, crop revenue will be measured as self-reported price per kg times quantity sold, and total profit will be measured both through a direct elicitation and by subtracting total reported costs from total revenues.

Another group of outcomes sheds light on the mechanisms through which water conservation may be affected: through adopting efficient irrigation technology, shifting to less water-intensive cropping patterns, or simply through irrigating less. To measure irrigation technology adoption, we will collect self-reported data on technological water conservation measures taken (such as micro-irrigation, alternate furrow irrigation, or mulching). We then create two technological water conservation outcomes: a dummy variable for whether any measure was taken, and an index of z-scores following Kling et al. (2007) where we set less intensive measures to 1 if a more intensive but mutually exclusive measure was taken (e.g. if alternate furrow is used, we set furrow to 1, since alternate furrow is a more intensive conservation strategy than furrow alone and the two are mutually exclusive). To measure changes in the water-intensity of planted crops, we both measure gross cropped area, and we will create an index of crop water requirements using data on cropped area and crop choice, parameterized by agronomic estimates of optimal water application rates. To understand the margins over which individuals adjust irrigation, we will measure irrigated area, irrigation frequency, and irrigation intensity.

A third group of outcomes investigates the possibility that conservation credits, as implemented on only one well, could cause farmers to substitute to other wells or water sources – a form of what is known as "leakage" in the PES literature. We will assess these substitution patterns with five outcomes: an indicator for the use of any other non-metered irrigation source, the area irrigated from other ground and surface water sources, and an estimate of irrigation water volume drawn from other ground and surface water sources (derived by multiplying together irrigated area, irrigation frequency, and irrigation intensity).

5 Analysis Plan

After checking for balance between treatment groups and attrition status, our analysis will proceed in three steps. First, we will report evidence on how individuals respond to groundwater prices through intent-to-treat (ITT) analysis of the conservation credits intervention as a whole. Second, we will estimate a model of demand for groundwater irrigation, using the price variation induced by our experiment in an instrumental variables strategy. Third, we will analyze the cost-effectiveness of the intervention from the perspective of a budget-constrained electric utility, by estimating the cost per unit of energy conserved.

5.1 Balance checks

Treatment Groups We will report baseline characteristics of the experimental sample in each treatment group including, but not necessarily limited to: household size, age of primary agricultural decision-maker, farm size (hectares), primary Rabi crop (including long-duration cotton planted before Rabi onset), irrigation technology, irrigation intensity, the power of the metered pumpset, and the depth-to-water and total dissolved solids in the metered well. We will test for balance using Wald test for the hypothesis that there is no difference between any of these variables in treatment and control groups.

Attrition. We will next test for potential differential attrition in treatment and control in each of our hours-or-irrigation and endline datasets. First, we will assess whether missing observations in each dataset are equally prevalent in treatment and control groups. Second, we will check for balance in missing and non-missing data using the same baseline characteristics as in our treatment balance checks. In the case that we find evidence of differential attrition of either kind, we will report bounds for all treatment effects following Lee (2009).

5.2 Program evaluation via ITT estimates

Primary statistical model. We will first report intent-to-treat (ITT) estimates of the effects of the conservation credits intervention as a whole, regardless of the specific sub-treatments. These estimates can be interpreted as a reduced-form measure of whether individuals respond to water prices. Each outcome variable will be compared between treatment (all farmers receiving any form of conservation credits) and control (farmers not receiving conservation credits). We will use ordinary least squares to estimate a monthly panel regression of the following form:

$$Y_{it} \quad = \alpha + \beta_1(\text{Conservation Credits})_i + X_{it}'\gamma + \mu_t + \varepsilon_{it}, \tag{3}$$

where Y_{it} is an outcome variable for farmer i in month t, $(\text{Conservation Credits})_i$ is an indicator for being in one of the conservation treatment groups, μ_t are month fixed effects, and X_{it} is a vector of individual-specific covariates (with time subscripts because some will be interacted with month indicators).

Covariates will include stratification variables (village indicators and an indicator for being above or below median forecasted hours of irrigation) interacted with month indicators. Covariates will also include baseline characteristics; to avoid overfitting and cherry-picking, we will use the double-LASSO method (following Belloni et al., 2013) to choose covariates from a high-dimensional set of variables derived from the baseline survey and other pre-randomization data. As a secondary specification, results will also be shown without these baseline characteristics. In all models, we will use cluster-robust standard errors, clustering by individual, to correct for arbitrary correlation in outcomes within individual across months. To increase power, we will exclude any month in which less than 50% of the control group has nonzero, non-missing observations – we try to schedule our program in months where pumping is the norm, but an unusual rainfall pattern may lead to unexpectedly low pumping in program months.

Primary outcome variable. The primary outcome variable will be hours of pump operation in each month of meter reading. Because the true functional form of the treatment effect is unknown, we will consider three specifications of this outcome variable. The first two are the untransformed hours of operation and its inverse hyperbolic sine (Asinh) transformation. These transformations will estimate the treatment effect more precisely if it is, respectively, constant (i.e., everyone reduces by an equal number of hours) or proportional (i.e., everyone reduces by an equal percentage of their hours absent the intervention). Regression coefficients from an Asinh transformation can be interpreted similarly as a natural log transformation (i.e., as proportional changes). We choose the Asinh over the natural log because the Asinh admits zeroes, unlike the natural log, albeit in a particular functional form.

We will also consider a third specification in which the dependent variable is the natural log of total hours of pump operation across all months of potential payments within each year, and farmers who never pump at all (likely a small number) are excluded. This specification models a proportional treatment effect but reduces the influence of the functional form choice for handling individual monthly zeros, by combining zero and nonzero observations into a single total. The downside is that this specification may cost some power; the time period is interpreted as one year instead of one month, so covariates cannot vary by month.

We are interested in both whether the program had an effect and the quantitative magnitude of the effect. To answer the first question, we will conduct one-sided t-tests ($\alpha = 0.05$), in which the alternative hypothesis (H_a) is that the program had a negative effect on pumping duration, and the null hypothesis (H_0) is that the program had a zero or positive effect. Because having three versions of the primary outcome constitutes multiple hypothesis testing, we will adjust the p-values using the free step-down approach of Westfall and Young (1993) following Kling et al. (2007).

To reduce the influence of outliers, final variables (after any transformations) will be winsorized by replacing extreme outliers with the next most extreme value. We define extreme outliers as values exceeding the third quartile plus three times the interquartile range, of the nonzero values of the same variable.

Secondary statistical models. In addition to the basic linear regression specification, we will show results from three other models. First, to investigate seasonal patterns in treatment effects, we will augment the primary regression to include time-varying (i.e., month-specific) treatment effects. Second, to more explicitly distinguish between the intensive and extensive margin of irrigation, we will apply a dynamic unobserved effects Tobit Type II model for censored values. Third, to reduce reliance on functional form, we will study the treatment effects across the full distribution using quantile regressions. We will simultaneously estimate quantile treatment effects at the 19 quantiles $\{0.05, 0.10, ..., 0.95\}$, obtaining standard errors via bootstrap to account for correlation across quantiles, using the following form:

$$Q_{Y_{it}}(\tau) + \beta_1(\text{Conservation Credits})_i + X_{it}'\gamma + \mu_t + \varepsilon_{it} \tag{4}$$

Secondary outcome variables. In addition to the primary outcome variable, we will use Equation 3 to examine the effects of the intervention on several other outcomes. These will enable us to better understand (1) the impacts as measured in units of energy and water, (2) whether the intervention has any measurable environmental or economic impacts, (3) the particular mechanisms through which farmers reduce water consumption, and (4) whether the intervention induced substitution ("leakage") to other water sources. The precise variables follow, with details of construction in Section 4.2:

1. Unit conversions: Implied energy consumption; implied water consumption (for each: monthly, monthly Asinh, and natural log of yearly totals).
2. Environmental impacts: Depth to water level; total dissolved solids.
3. Economic impacts: Crop revenue; farm profits (for each: level, Asinh, and both per hectare).
4. Mechanisms: Water conservation measures; gross cropped area; crop water intensity; gross irrigated area; irrigation frequency; irrigation intensity.
5. Leakage: Use of other irrigation sources; gross area irrigated by other sources, water volume used for irrigation from other sources.

Because secondary outcomes will be measured annually following each Rabi season, in these regressions, the time period in Equation 3 will be interpreted as one year instead of one month. For the outcomes in the unit conversion and economic impact categories, we will adjust p-values category-wise using the same method as for the primary outcomes. For the outcomes in the environmental impact, mechanism, and leakage categories, we plan not to adjust the p-values, because each outcome answers a different question, and they are not measures by which we will judge the overall success of the intervention. All variables will be winsorized in the same way as the primary variables.

Heterogeneity. We also will analyze treatment effect heterogeneity according to baseline characteristics in order to further assess mechanisms of conservation and the feasibility of better targeting the program. These analyses will involve a variation of the regression in Equation 3, in which the treatment indicator is interacted with an exhaustive set of indicators in a particular category. Three of these categories will be: (1) whether the farmer is a medium or large landowner (defined as owning more than 2 hectares of land), (2) whether the farmer had previously invested in micro-irrigation technology such as sprinkler or drip irrigation, and (3) whether the farmer shares their primary irrigation source with others.

Two additional heterogeneity analyses will serve as indirect tests of leakage. The idea of these tests is to check whether the treatment effect is larger in areas where there are more opportunities for leakage to other groundwater sources. Specifically, we would expect to see more leakage in areas where more un-priced groundwater is available. In these tests, we will interact the treatment indicator with (1) an indicator for whether the farmer has unmetered groundwater sources (a proxy for the availability of on-farm unpriced groundwater), and (2) the number of neighboring farms that have a groundwater source but are not enrolled in the conservation credit treatment (a proxy for the availability of off-farm unpriced groundwater). In both cases, a positive interaction coefficient is evidence of leakage. Unlike the direct tests, which compare water use from secondary groundwater sources among treated and untreated farmers, the indirect tests have the advantage that they do not rely on self-reported data.

5.3 Demand estimation via instrumental variables (IV)

Our second analysis will estimate a model of demand for groundwater irrigation. We will use a linear regression to predict duration of pump operation:

$$Y_{it} = \alpha + \beta p_{it} + X_{it}'\gamma + \mu_t + \varepsilon_{it} \tag{5}$$

where $p_{it} \in \{0, 20, 40\}$ indicates the marginal cost of an hour of irrigation for farmer i in month t, and μ_t are month fixed effects. Again, X_{it} is a vector of individual-specific covariates including stratification variables interacted with month indicators, plus baseline characteristics chosen by double-LASSO. Standard errors will be clustered by individual.

We will estimate Equation 5 by two-stage least squares to correct for endogeneity in price. Note that while Control farmers always face a price of 0, Conservation Credits farmers in the *low price* sub-treatments face a price of either 0 or 20, and those in *high price* sub-treatments face a price of either 0 or 40, depending on whether their consumption is above or below their benchmark. This introduces endogeneity into Equation 5: in the Conservation Credit treatment, positive consumption shocks ε_{it} are mechanically correlated with zero prices, biasing OLS estimates of β downward.

To boost precision in this estimate while avoiding overfitting and weak instruments concerns, we will apply the instrumental variables LASSO method of Belloni et al. (2012). Our set of candidate instruments will consist of indicators for each of the four conservation credit sub-treatments, and their interactions with baseline characteristics. The final set of instruments will be chosen from this candidate set by the algorithm. As a secondary specification, we will also show results using only the four sub-treatment indicators as instruments.

The IV estimate of β can be interpreted as a local average treatment effect of our experimental price variation on those farmers whose marginal consumption is priced. Our methodology is in the spirit of quasi-experimental estimates of the elasticity of taxable income from non-linear budget sets (as summarized by Saez et al., 2012) and of electricity demand (Ito, 2014).

The exclusion restriction for these instruments is that the Conservation Credit sub-treatment does not affect consumption except through the actual price of irrigation. This assumption will be violated if the Conservation Credit sub-treatments affect consumption even for farmers who do not face positive marginal incentives in a given month – for example, if they attempt to conserve below the benchmark but fail to reach their target. This is one limitation of our empirical strategy.

Outcome variables. The primary outcome variable will be hours of pump operation as measured in meter readings. We will consider the same three specifications as in the intent-to-treat analysis, adjusting p-values in the same way: (1) the untransformed measure in each month, (2) an inverse hyperbolic sine (Asinh) transformation, and (3) the natural log of the yearly total hours. Secondary outcome variables will be the unit conversions: implied water consumption, and implied energy consumption.

Heterogeneity analysis. We will again explore the first three dimensions of heterogeneity for demand as for intent-to-treat effects: farm size, micro-irrigation technology, and well sharing.

5.4 Cost-effectiveness

Our third analysis will consider the cost-effectiveness of the conservation credits intervention as implemented in the study. Because groundwater conservation yields the side benefit of reduced electricity demand, a conservation credits program could be implemented by a budget-constrained electric utility under one condition: that the cost of the energy conserved is larger than the cost of the program. We explore the viability of this idea through three questions. For each, we will report answers for the program as a whole, as well as the *low price* treatment group alone (discarding data from the *high price* treatment group).

Question 1. *What is the minimum marginal cost of electricity for which the program could be implemented by an electric utility with a budget-balance constraint?*

For each unit of energy conserved through the program, an implementing electric utility reduces its costs by the marginal cost of electricity. Therefore, the minimum marginal cost (for which the program is viable) is equal to the average cost of the program per unit energy conserved. This will be calculated using the following formula:

$$\text{Cost per unit energy conserved} = \frac{\text{Total monthly expenditures}}{\text{Total monthly energy conserved}} \tag{6}$$

Total monthly expenditures will be tabulated from program data. To estimate total monthly energy conserved, we will multiply the linear treatment effect of the program on energy consumption (from the intent-to-treat analysis) by the number of participants in the treatment group. To obtain a confidence interval that allows for correlation between the numerator and denominator, we will bootstrap the treatment effect and expenditures together, stratifying resampling draws on treatment assignment. Note that a limitation of this confidence interval is that it will ignore uncertainty in pump efficiencies.

Question 2. *At the best estimate of actual marginal costs faced by electric utilities in India, could the program be implemented by an electric utility with a budget-balance constraint?*

To answer this question, we will conduct a literature review on the marginal costs of electricity provision in India and arrive at a best estimate of the typical marginal cost prior to performing these calculations. Then, we will conduct a one-tailed bootstrap test ($\alpha = 0.05$), where the null hypothesis that the cost per unit energy conserved is greater than or equal to this marginal cost, while the alternative hypothesis is that the cost per unit energy conserved is less than this marginal cost (revenue-positive). Ignoring uncertainty in the marginal cost estimate, we will draw 1,000 bootstrap samples and count the number in which the cost per unit energy conserved exceeds the marginal cost of electricity. The null hypothesis is rejected if this condition is met for fewer than 950 draws (95 percent).

Question 3. *What is the minimum subsidy per unit of energy that an electric utility with a budget-balance constraint would require to implement the program?*

We will consider this question only if the answer to Question 2 is no (i.e., we fail to reject the null hypothesis). Even if a conservation credits program does not pay for itself through electricity cost savings, a government placing social value on groundwater conservation may be willing to subsidize the program. We will calculate the minimum necessary subsidy by subtracting the estimated cost of the program per unit energy conserved from the best estimate of the marginal cost of electricity. Again ignoring uncertainty in the marginal cost estimate, the confidence interval will be calculated in the same way as in Question 1.

6 Pilot Results

We conducted a small pilot of this experiment among 90 farmers in three villages in a nearby district of the same state (Khambhalia, Gujarat). This pilot informed the experimental design in four ways.

Demonstrates logistical feasibility. The pilot shows that the intervention can be successfully implemented among a similar population as the experimental sample. First, farmers were broadly willing to participate, voluntarily accept hours-of-use meters and agree to monthly meter readings. Of 144 farmers randomly sampled from village rosters, 100% agreed to allow us to install a meter on their pump. Of 90 farmers meeting eligibility criteria, meters were successfully installed for 100%. Of the same group, one withdrew during the intervention, yielding a 99% completion rate.

Second, meter tampering is difficult and appeared to be minimal. The meter itself is sealed, with no controls other than a reset button (which can be easily detected after a first reading). Disconnection is not simple and leaves indications in the form of uncoiled wires; only two farmers showed evidence of having disconnected and reconnected in the same month. Third, farmers appear to understand the program; during the initial intervention visit, farmers were asked questions designed to measure comprehension and corrected if necessary; surveyors reported a subjective assessment that most farmers understood the program very well.

Improvements in intervention design. The pilot yielded several ideas for improving the effectiveness and power of the intervention that we will incorporate into the experiment. First, 20 percent of farmers permanently disconnected their meters following their last irrigation of the season. However, disconnecting the meter disqualified treatment group farmers from receiving payments, so the disconnections were highly concentrated in the control group (14 of 18). To ensure accurate data from the control group, all farmers will be offered a small financial reward to keep their meter connected through the end of the meter-reading period.

Second, the experiment will focus on months and geographical regions in which the vast majority of farmers have access to groundwater (i.e., without deepening a well). In the pilot, 29 percent of meter readings showed zero consumption, a pattern that rose to 50 percent by the end of the pilot. Discussion with farmers revealed that many had stopped pumping because their well had gone dry. These zeros substantially reduced statistical power (by increasing the variance of the outcome variable), and paying farmers whose well had gone dry was perceived to be unfair by the implementing partner. For the experiment, the geographical region was chosen in part because it has more reliable water availability. In addition, conservation credits will be paid during a more limited number of months (those in which a large majority of farmers are known to irrigate crops).

Sample size calculations. Neither water consumption nor our proxy, hours of pump operation, is often measured at the farm level in India, and so our pilot measurements represent a contribution in themselves. Figure 5 plots the full distribution of monthly measurements of pump operation time across all farmers in our pilot. The variance is much larger than initially expected, which informed our power calculations and led us to revise upward the sample size of the experiment.

Suggests conservation credits may yield the expected effects. While the pilot has low statistical power and cannot yield precise results, analysis is not inconsistent with the intervention having the expected direction of response and a large effect magnitude. Figure 6 plots the mean number of hours pumped per month of the pilot. Before the price incentive was introduced, farmers in the treatment group pumped for more mean hours than those in control; after conservation credits began, the treatment group pumped for fewer mean hours than control each month (although none of these differences are statistically significant).

To quantify these differences, Table 1 shows the results of linear regressions following Equation 3, with each column including a different set of covariates. Point estimates suggest that eligibility for conservation credits induces a practically large reduction in pumping hours: a 32 percent decrease on a control-group mean of 38 hours per month. Although these point estimates are imprecise (confidence intervals include both zero and some positive values), they appear to be stable across specifications.

7 Conclusion

This paper presents an experimental protocol to estimate the demand response to groundwater pricing in irrigated agriculture in a region of Gujarat, India. To estimate demand, the study will introduce random variation in prices through an intervention that offers payments for groundwater conservation relative to a benchmark quantity. We show how to use our demand estimate - given a marginal social damages function - to calculate the optimal Pigouvian groundwater tax in our setting. The optimal quantity of groundwater conservation could be achieved through a marginal incentive on either agricultural electricity, groundwater, or duration of pump operation.

Our study will also evaluate the effectiveness of conservation credits, a second-best policy solution similar to "payments for environmental services" programs. In many settings, Pigouvian taxes may be politically infeasible. By exchanging corrective subsidies for taxes, this program overcomes the political barriers to taxing the agricultural sector, while still introducing marginal incentives for conservation. This may be a promising policy approach for reducing inefficient groundwater extraction.

Conservation credits, however, are generally not efficient, unless benchmarks can be perfectly targeted or revenue constraints do not bind. Some farmers are likely to be extra-marginal: their extraction is so far beyond their benchmark that they do not benefit from conservation. This raises another important question: what is the optimal conservation credit program that a donor or government would be willing to implement? Evaluating this question, given the goals and constraints of a funder, depends not only on aggregate groundwater demand, but also on the distribution of utilization across farmers. Future research may be able to use variation in the program design parameters, like that introduced in this study, to make progress on this question.

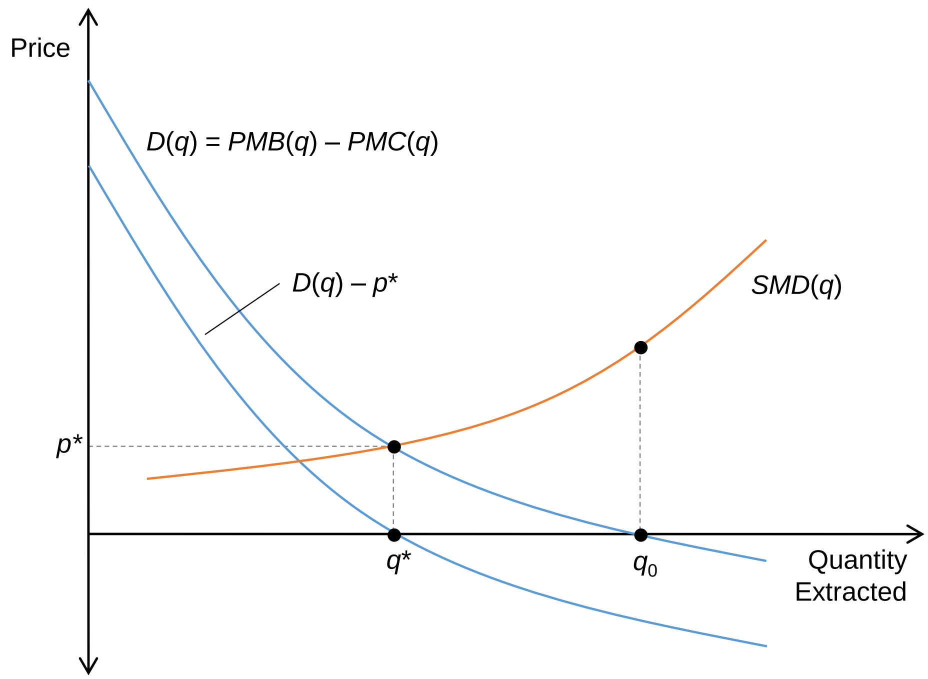

Figure 1: Price regulation for groundwater management.

This figure shows how price regulation can be used to achieve the optimal groundwater quantity extracted. Inverse demand for groundwater $D(q)$ is the difference between private marginal benefits $PMB(q)$ and private marginal costs $PMC(q)$; groundwater extraction also creates social marginal damages $SMD(q)$. Without regulation (i.e., at a price of zero), irrigators will consume the amount where demand meets the x-axis, q_0. When the price is set to p^*, the value of social marginal damages when it equals demand, irrigators will internalize the social damages, shifting effective demand down such that they instead consume the optimal quantity q^*.

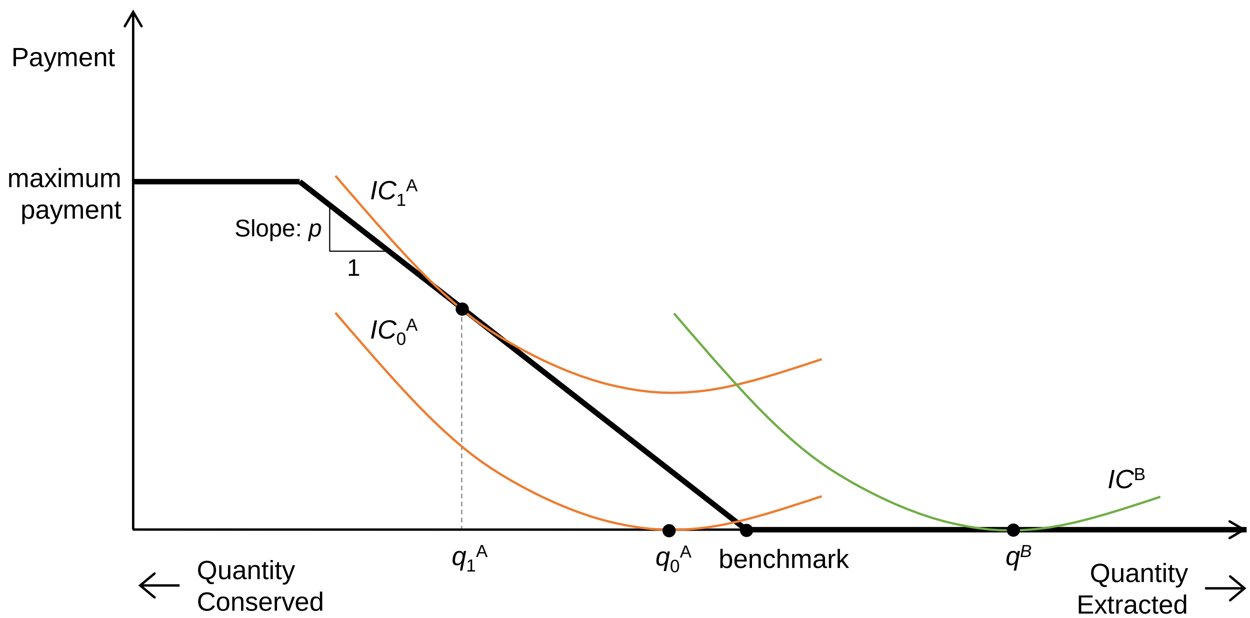

Figure 2: Budget set of conservation credits.

This figure shows the general form of the budget set created by a conservation credit program, along with indifference curves of two representative participants. The payment equals the price p times the quantity units conserved below the benchmark, up to a maximum payment. Irrigator A is marginal and will respond to the program by reducing quantity extracted. Irrigator B is extra-marginal, and does not change quantity extraction in response to the program.

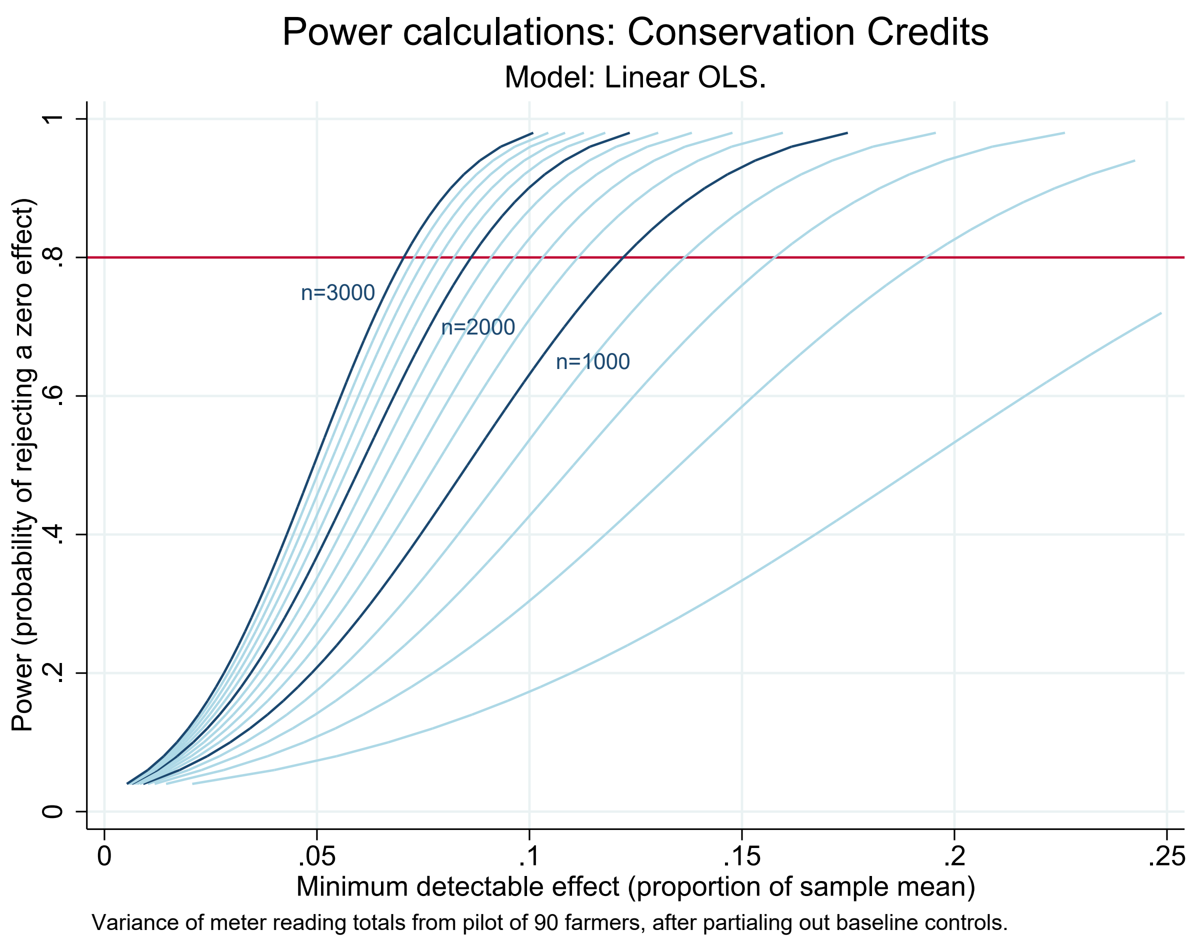

Figure 3: Power curves for a range of sample sizes and minimum detectable effects.

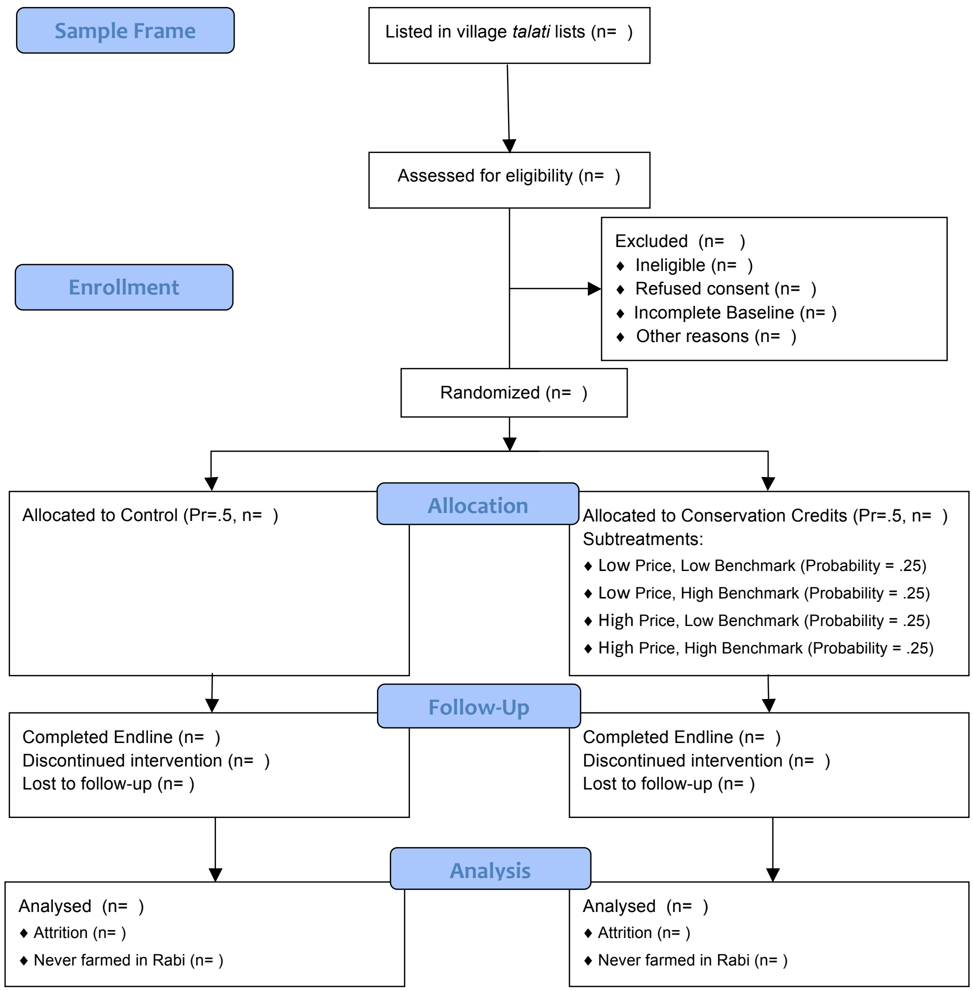

Figure 4: Experimental Design

Notes: This figure displays the design of a randomized experiment to estimate the demand response to agricultural groundwater prices.

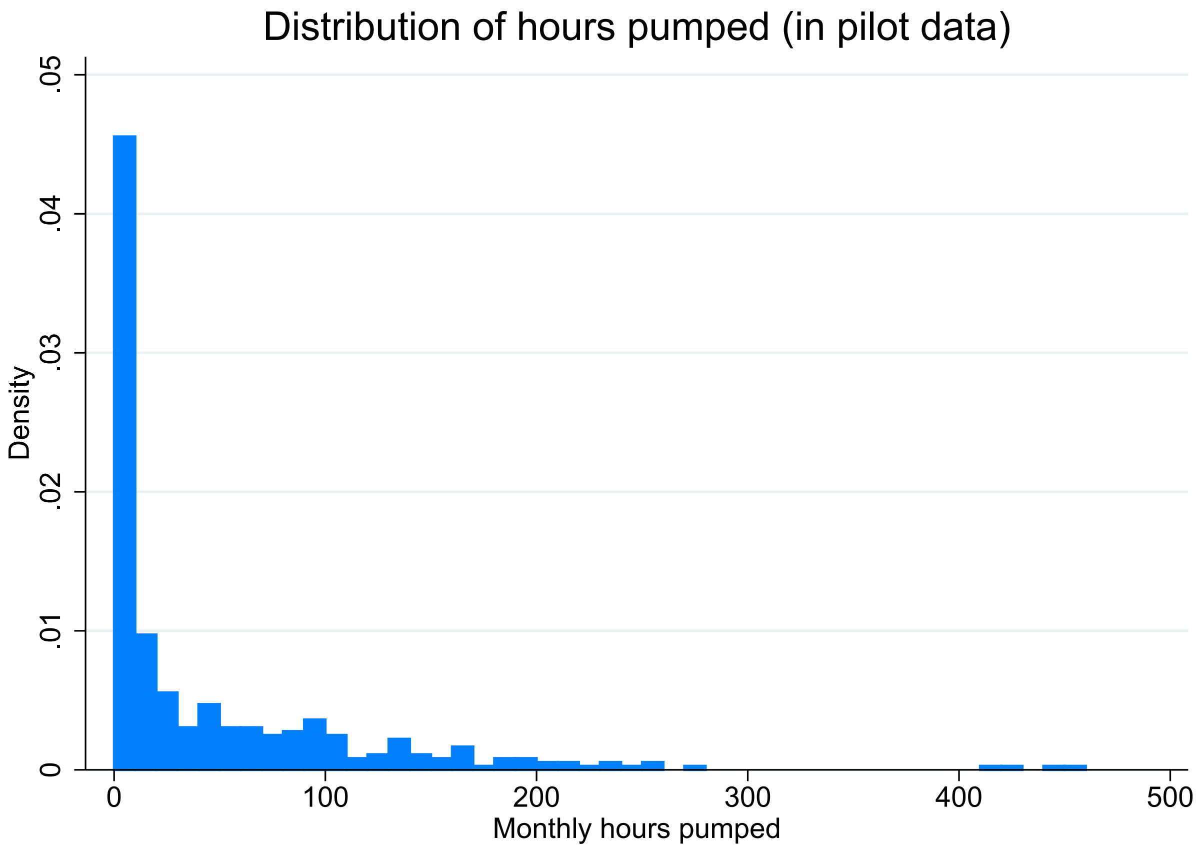

Figure 5: Distribution of monthly hours of groundwater irrigation, pooled across months

Notes: This figure plots the histogram of the monthly hours of groundwater irrigation measured in the 2017-2018 Rabi season (October-February) in our pilot study of 90 farmers in Khambaliya, Gujarat.

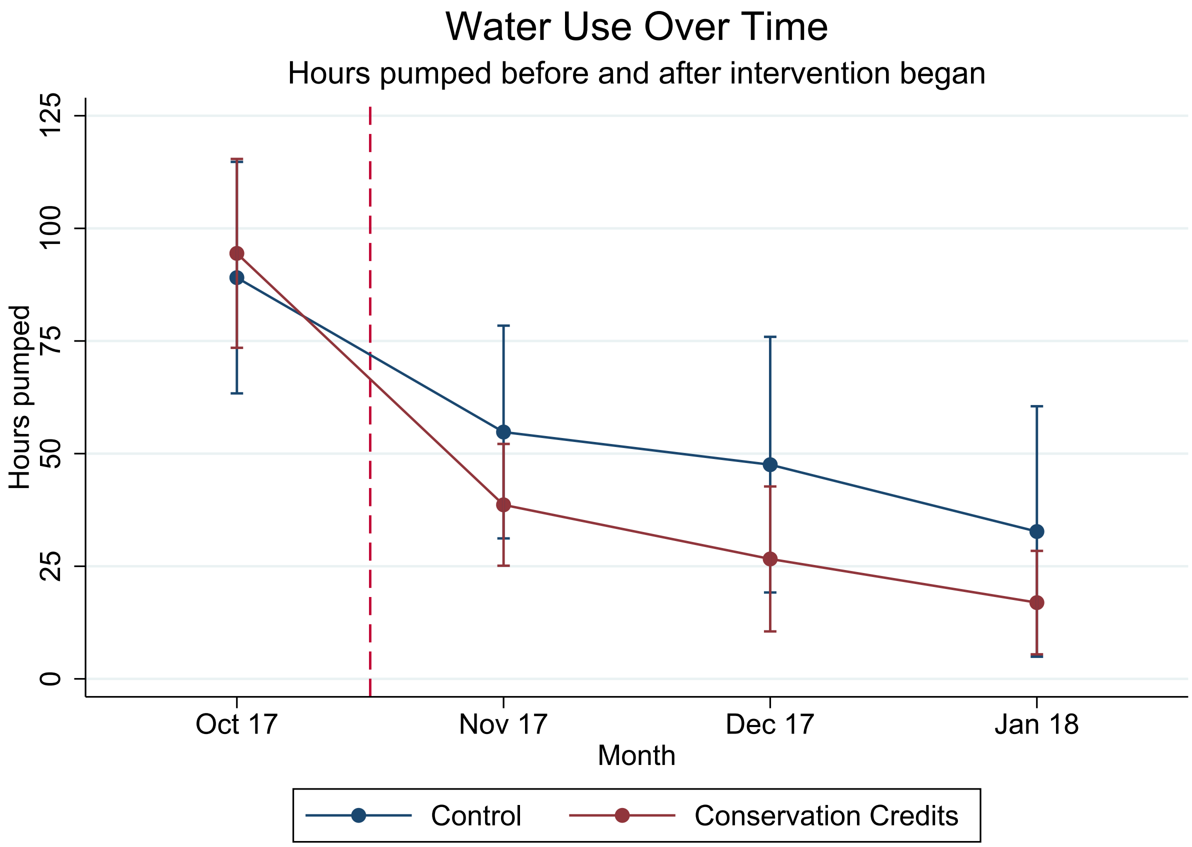

Figure 6: Event Study

Notes: This figure plots the average monthly hours of groundwater irrigation among farmers in our pilot experiment over the winter 2017-2018 program. The bars denote the standard errors of the mean.

	Monthly Pumping Hours				
	(1)	(2)	(3)	(4)	(5)
Conservation Credit treatment group	-12.067	-12.126	-16.666	-12.032	-16.872
	(12.071)	(11.423)	(13.305)	(15.391)	(15.672)
Strata FE		X	X	X	X
Month FE			X		X
Sub-village FE			X		X
Baseline controls				X	X
Observations	270	270	270	258	258
Clusters	90	90	90	86	86

Standard errors clustered by individual.

Table 1: ITT effect of conservation credits program in pilot.

References

Antmann, P. (2009). Reducing technical and non-technical losses in the power sector. Technical Report July 2009.

Badiani, R. and Jessoe, K. (2017). Electricity Prices, Groundwater and Agriculture: The Environmental and Agricultural Impacts of Electricity Subsidies in India. In Schlenker, W., editor, *Understanding Productivity Growth in Agriculture*. University of Chicago Press.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain. *Econometrica*, 80(6):2369–2429.

Belloni, A., Chernozhukov, V., and Hansen, C. (2013). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650.

Börner, J., Baylis, K., Corbera, E., Ezzine-de Blas, D., Honey-Rosés, J., Persson, U. M., and Wunder, S. (2017). The Effectiveness of Payments for Environmental Services. *World Development*, 96:359–374.

Bruno, E. M. (2018). Agricultural Groundwater Markets: Understanding the Gains from Trade and the Role of Market Power. *Job Market Paper*, pages 1–54.

Central Ground Water Board (2013). Ground Water Year Book - India 2012-13. Technical report, Ministry of Water Resources, Government of India, Faridabad.

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

Golden, M. and Min, B. (2012). Corruption and theft of electricity in an Indian state.

Gonzalez-Alvarez, Y., Keeler, A. G., and Mullen, J. D. (2006). Farm-level irrigation and the marginal cost of water use: Evidence from Georgia. *Journal of Environmental Management*, 80(4):311–317.

Hendricks, N. P. and Peterson, J. M. (2012). Fixed Effects Estimation of the Intensive and Extensive Margins of Irrigation Water Demand.

Ito, K. (2014). Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. *American Economic Review*, 104(2):537–563.

Jack, B. K. and Cardona Santos, E. (2017). The leakage and livelihood impacts of PES contracts: A targeting experiment in Malawi. *Land Use Policy*, 63:645–658.

Jack, B. K. and Smith, G. (2016). Charging Ahead: Prepaid Electricity Metering in South Africa. *NBER Working Paper Series*.

Jayachandran, S., De Laat, J., Lambin, E. F., Stanton, C. Y., Audy, R., and Thomas, N. E. (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science*, 357(6348):267–273.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1):83–119.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3):1071–1102.

Lee, K., Miguel, E., and Wolfram, C. (2018). Experimental Evidence on the Economics of Rural Electrification. *Working Paper*.

McRae, S. (2015). Infrastructure Quality and the Subsidy Trap. *American Economic Review*, 105(1):35–66.

Meenakshi, J., Banerji, A., Mukherji, A., and Gupta, A. (2013). Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India. *3ie Impact Evaluation Report*, 4.

Ministry of Agriculture of the Government of India (2014). Agriculture Census 2010-11.

Northeast Group LLC (2014). Emerging Markets Smart Grid: Outlook 2015. Technical report.

Pattanayak, S. K., Wunder, S., and Ferraro, P. J. (2010). Show me the money: Do payments supply environmental services in developing countries? *Review of Environmental Economics and Policy*, 4(2):254–274.

Registrar General and Census Commissioner of India (2001). Census of India 2001.

Registrar General and Census Commissioner of India (2011). Census of India 2011.

Saez, E., Slemrod, J., and Giertz, S. H. (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, 50(1):3–50.

Samadhan E Cube Innovator Pvt. Ltd. (2016). Baseline Study for project villages of Coastal Salinity Prevention Cell in Gujarat under Kharash Vistarotthan Yojana (KVY): Talaja and Gogha Taluka, Bhavnagar District. Technical report, Coastal Salinity Prevention Cell.

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.

Sovacool, B. K. (2017). Reviewing, Reforming, and Rethinking Global Energy Subsidies: Towards a Political Economy Research Agenda. *Ecological Economics*, 135:150–163.

Westfall, P. H. and Young, S. S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. Wiley, New York.