Online Appendix for:

Rationing the Commons

Nicholas Ryan and Anant Sudarshan

This Online Appendix contains supplementary materials for the above-referenced article. There are five lettered appendices. Appendix A discusses our data sources, with a particular focus on measures of groundwater depth and their relation to geological factors. Appendix B gives derivations in our model omitted from the text. Appendix C gives robustness checks for our estimates of the effect of depth on profits and other results. Appendix D gives additional empirical results beyond those in the main text. Finally, Appendix E presents a dynamic extension of our model and uses this extension to calculate the opportunity cost of water.

A Appendix: Data

This appendix describes our data sources and the construction of several important variables. Part A a shows that well depth is a strong proxy for groundwater depth. Part A b describes the theory and data for the prediction of groundwater based on subsurface geology. Part A c describes the calculation of water input.

a Relation between well depth and groundwater depth

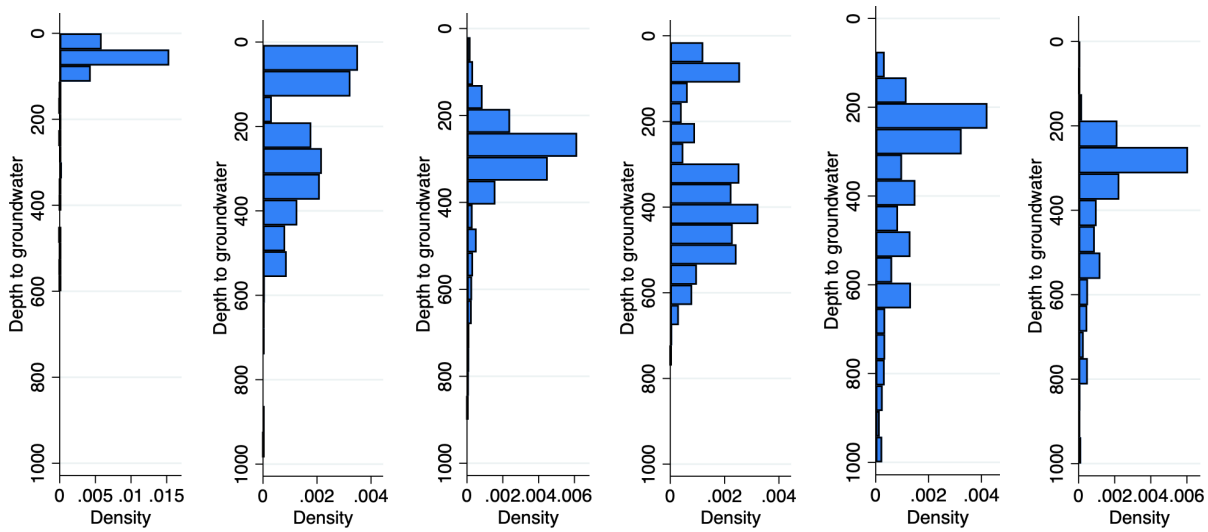
Farmer water extraction, as governed by equation 1, depends on how far down it is to groundwater. Our survey measures, instead, the depth of a farmer's well, since water levels fluctuate and farmers generally know their well depth better. This section uses ancillary data to study how well depth is related to groundwater depth. We find a tight, linear relationship between well depth and groundwater depth in two data sources, which justifies using well depth as a proxy for access to groundwater.

Well depths are closely related to groundwater levels, since if water is further down, a farmer has to bore the well deeper to reach it. However, well depth is not the same as water depth—in general, for active wells, the well will go down deeper than the water. This margin of extra depth

is kept because water may seep into a well only slowly after it is extracted, because water levels fluctuate from year to year, depending on the amount of monsoon rainfall, and because the average water depth is increasing over time, so farmers boring a well leave a margin of depth to account for future groundwater depletion.

Figure A1 shows the variation in well depth in our sample. Dug SDO, in panel A, has the shallowest water, with most wells less than a hundred feet deep, whereas in Mundawar many farmers have wells greater than four hundred feet deep. Even within SDOs there is a large dispersion in well depths. For example, in Nainwa (panel D), the mean well depth is around 300 feet but wells range from less than 200 to over 500 feet in depth.

Figure A1: Variation in Depth



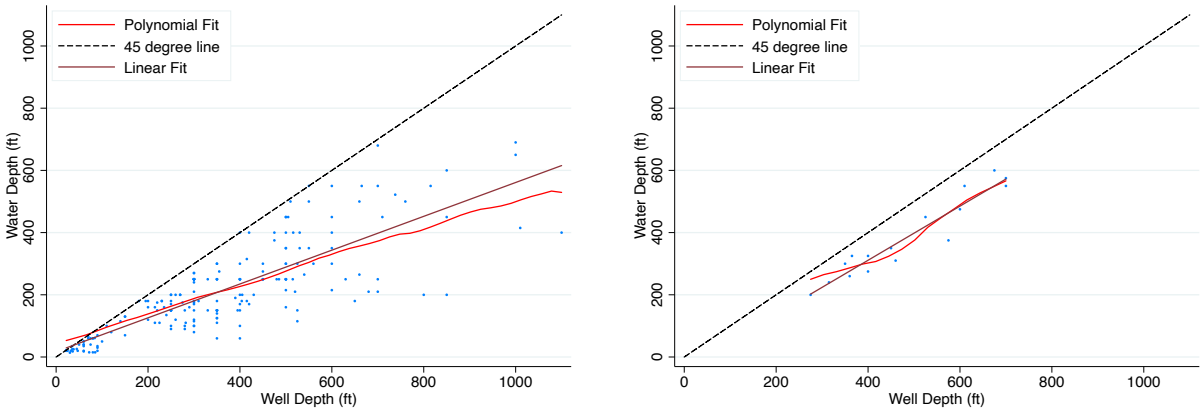
In order to compare well depth and water depth, we gathered two separate samples: a sample of farmers and a sample of the boremen who dig wells for a living.

Farmer phone survey. To get an estimate of the relationship between water depth and well depth, 200 farmers from the main survey sample were asked to estimate groundwater depth in their wells. Farmers were sampled from all the six SDOs used in the main analysis.

Bore drilling agents survey. Farmers may not know groundwater depth as well as professionals who drill for a living. We therefore contacted bore drilling agents operating in five SDOs from the main survey data, excluding Dug. In total, we contacted 20 bore drilling agents with approximately 4 agents in each SDO. Of these 20, 16 agents replied to our brief survey.

Figure A2 shows the results of these two surveys. Panel A, on the left, shows results from the farmer survey, and Panel B, on the right, from the bore drilling agent survey. In both of these samples, water depth is less than well depth, as we should expect from active wells—if the groundwater table is below the bottom of the well, the well is dry and would not be used. Bore drilling agents report that wells are dug 50 to 75 feet deeper than water levels. In both of these samples, we observe a tight, positive, linear relationship between well depth and water depth.

Figure A2: Water Depth versus Well Depth (in feet)



A. Farmer Phone Survey B. Bore-Drilling Agent Survey

The figure shows the relationship between well depth and water depth from the two sources that we have mentioned above. Both figures plot a linear and a polynomial fit along with the 45° line. The polynomial fit is done using lpoly function in Stata with a bandwith of 50. In the second figure, as we do not have any bore-drilling agent who responded to us from Dug, there is low representation of lower well depths.

Table A1 shows the coefficients from the linear regression lines plotted in Figure A2. The coefficients on well depth are positive and precisely estimated. For the borewell agent survey, moreover, we estimate a coefficient of 0.871 (standard error 0.0728), which is not significantly different than one. In this case, the relationship between well depth and groundwater depth is not only linear, but, with a coefficient of one, variation in depth from one measure is one-to-one with variation in the other. We therefore conclude that variation in well depth is an appropriate proxy for variation in groundwater depth in a farmer's well.

b Use of geology to predict groundwater conditions

A number of studies have shown the predictive power of geological features for groundwater availability (Sander, 2007; Jasmin and Mallikarjuna, 2011; Mallick et al., 2015). Geological factors

Table A1
Relationship between water depth and well depth

	(1) Farmer Phone Survey	(2) Bore Drilling Agent Survey
Well Depth (feet)	0.544***	0.871***
	(0.0278)	(0.0728)
p-value: Well-Depth = 1	0.00	0.10
Observations	199	16

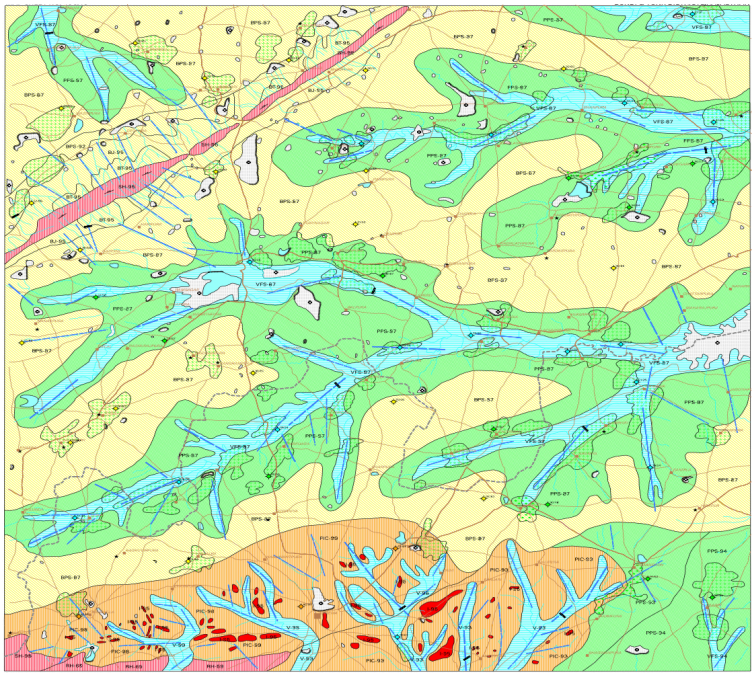
The table shows how well depth reported by the farmers in the survey relates to the actual water depth in the SDO that the farmer resides in. The first column shows results reported by the famers that were interviewd via phone. The second column shows results from the bore drilling agent survey. Finally in the third row, we report p-values to test if the coefficient on well depth is one. The statistical significance of a coefficient at ertain thresholds is indicated by * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

affecting groundwater include aquifer material or rock type, lineaments, geomophology, and topography. Mallick et al. (2015) create a groundwater potential index using ten layers, including geology and topography, to predict groundwater potential, and show that their predictions are correlated with true water depth and well flow rates. Sander (2007) and Lee, Park and Choi (2012) both find that lineaments, which we call fractures, are among the most important determinants of groundwater availability in hard rock aquifers. Blakeslee, Fishman and Srinivasan (2020) demonstrate, with direct field measurements in a subset of their sample in Karnataka, that farmer wells that intersect a greater number of fractures are more likely to have water.

To measure factors predictive of groundwater availability we obtained data from the Bhuvan Bhujal project (introduced in Section 2). The Ministry of Drinking Water and Sanitation, Government of India tasked the National Remote Sensing Centre, Hyderabad with producing groundwater prospects maps. The goal of the project was to find high yielding and sustainable borewell locations. The project produced a GIS database of geological features accurate down to a kilometer. Map layers can be viewed at the Bhuvan Bhujal Ground Water Prospects Information System: `https://bhuvan-app1.nrsc.gov.in/gwis/gwis.php`.

Figure A3 gives an example of a groundwater prospect map for Bundi district, Rajasthan. The colored areas are the type of rock underlying an area. The dashed blue straight lines indicate lineaments, underground fractures in rock that are conducive to the flow of water in hard rock aquifers. There is variation in the type of rock and aquifer, and the precise location of lineaments, down to a fine geographic scale. These factors underlie our predictions for groundwater depth.

Figure A3: Groundwater prospects map, Bundi district, Rajasthan



The figure shows a groundwater prospects map for the Bundi district of Rajasthan. The colored areas are the type of rock underlying an area. The dashed blue straight lines indicate lineaments, underground fractures in rock that are conducive to the flow of water in hard rock aquifers.

The groundwater prospectus maps include a rich set of features that the literature has identified as useful for the prediction of groundwater depth, including the type of rock; the porosity-permeability of a geological formation; faults, fractures and aquifers. Rock types are not simple but highly differentiated based on the porosity and schistosity (cleavage or fracturing) of the rock. Our instrument set uses only geological features, like the type of rock underground and fractures in that rock, which are plausibly excludable, since they do not directly affect surface productivity. We omit all surface features, such as topography, from our instrument set and instead include these features as exogenous controls.

c Calculation of water input

Water use by farmers is not metered, so it is necessary to calculate water use on the basis of pump use and groundwater conditions. Our survey was designed to ask farmers about variables that affect water extraction, including pump size, pump use and well depth. We use the survey variables to calculate water input in liters, for each farmer-crop, following a standard engineering formula for

water extraction (Manring, 2013). Water extraction is given by

$$W_i(H_i, D_i) = \rho \frac{P_i H_i}{D_i}.$$

A farmer with pump capacity P_i runs their pump for H_i hours per day to lift water from depth D_i. The physical constant ρ is given by

$$\rho = c\frac{E}{dg}$$

where c is a constant to correct units and account for friction, E is the pump efficiency, d is density of water, and g is the gravitational constant.

Table A2 gives the values of all the constants used in calculating water input. Our survey elicits all of the other variables that enter the water extraction function. The mean water input by farmer-crop, calculated in this manner, is roughly 1.5 million liters per season (see Table 1).

Table A2
Constants used in water input calculation

Variable	Value	Units
c	3.6×10^6	
E	0.25	
d	10^3	kg/m^3
g	9.8	m/s^2

The table shows the values of the constants used in the construction of water input. The density of water d and gravitational constant g are standard (Manring, 2013). The constants c and E are from studies of irrigation pumping in India (Shakti Foundation, 2016; Oxford University Press, 2011).

d Soil quality controls

Our survey did not collect soil quality measures since taking and analyzing soil samples is costly. We use village-level data on soil quality collected by the Indian government as part of the Soil Health Card Scheme. Launched in February, 2015, the scheme aims to provide farmers in India with cards that document the quality of the soil and recommendations to improve soil health.

Our data consists of categorical measures of soil nutrition status aggregated to the village level. The soil health variables are categorical measures of acidity/alkalinity and concentrations of phosphorus, potassium, copper, iron and zinc. We observe the total number of farmers in each

village that fell within certain ranges for each parameter, e.g. the number of farmers with highly alkaline soil. We transform the number of farmers within each group into proportions and then merge the soil quality dataset into the farmer survey by matching on district, block and village names. We match 49% of sample villages by exact name and 77% of sample villages including approximate name matches.

e Weather during the rabi season

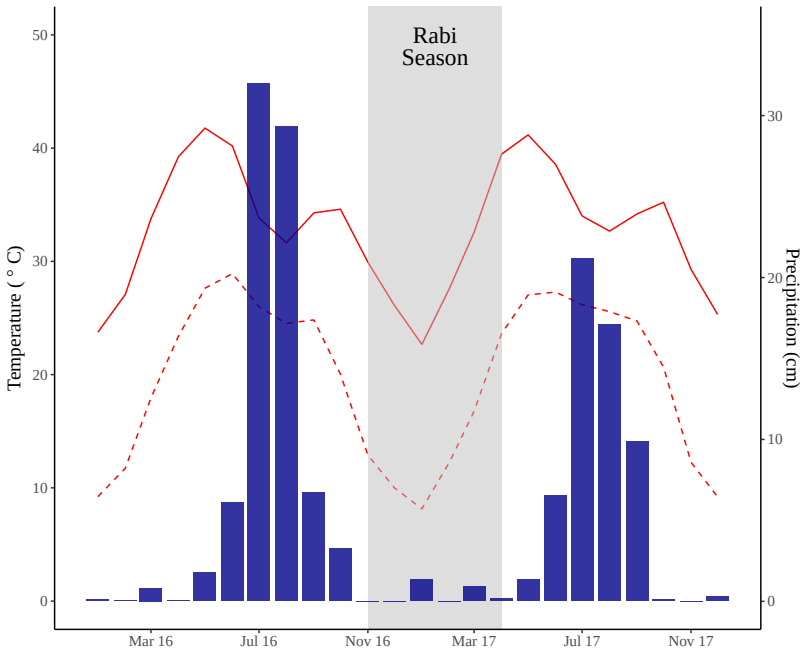
Our main specifications do not include controls for weather, though weather is an exogenous determinant of productivity. We do not control for weather is because there is no usable data that captures variation in weather at the required spatial scale within a single season. This subsection describes the climate during the Rabi season and discusses possible sources of weather data.

Our survey data come from the Rabi season of 2016-17. The Rabi season, with planting in late October or November and harvest in April, typically has minimal rainfall. Figure A4 shows rainfall and the minimum and maximum temperature during the season our data referenced. Each of these variables is averaged over all of the SDOs that comprise our survey sample. There is minimal rainfall during the referenced cropping season, several centimeters, and most of that is in early October, when most farmers have not yet planted. The pattern during and around our sample period is typical of the Hot Semi Arid (BSh) climate in Rajasthan, with rainfall highly concentrated in the preceding monsoon.

We investigated using cross-sectional information on precipitation or temperature as controls. The main sources of data are based on model imputations and, despite high nominal resolutions, do not have adequate true resolutions to control for weather variation within an SDO-season. An SDO is approximately 500 km^2 hence 20 km on a side. The TerraClimate dataset includes precipitation in grid cells of 4 km resolution; however, since there are far fewer rainfall stations, most of the data is imputed, and we found virtually no spatial variation in rainfall conditional on SDO fixed effects.

The only temperature data that has adequate nominal resolution is model-imputed rather than being collected by direct observation. The MODIS Land Surface Temperature and Emissivity data includes land surface temperature in grid cells of 1 km resolution. This data shows little variation in temperature across space within an SDO. We include temperature as a control in some specifications (Table C6). We omit these temperature variables from our main specification

Figure A4: Weather during Rabi season 2016-2017



The figure shows the weather during and around the cropping season that our data reference. The solid bars show the precipitation each month, measured against the right axis. The solid line shows the average daily maximum temperature each month and the dotted line the average daily minimum temperature, measured against the left axis. The grey shaded region gives approximately the duration of the Rabi cropping season.

because the MODIS model for land surface temperature includes ground cover as a covariate in a predictive model of surface temperature. For example, heavily forested or agricultural areas are assigned lower surface temperatures. This model-based imputation is problematic, since it would imply that temperature, which is meant as an exogenous control, would be endogenous to surface characteristics and agricultural productivity in particular.

B Appendix: Model

a Derivation of efficiency loss under rationing

Using the notation of Section 3 C, where production is only a function of hours of power use, the mean surplus levels under a Pigouvian regime and under rationing, respectively, are

$$\begin{aligned} S^P &= \mathbb{E}[\Omega_i \hat{F}(H_i^*)] - \mathbb{E}[H_i^*]p_H^* \\ S^R &= \mathbb{E}[\Omega_i \hat{F}(\overline{H})] - \mathbb{E}[\overline{H}]p_H^* \end{aligned}$$

where H_i^* is each farmer's chosen optimal level of pump use, in the Pigouvian regime, and $\overline{H}$ is the uniform level of pump use in the rationing regime. The difference between mean surplus in the two regimes is

$$\begin{aligned} S^P - S^R &= \mathbb{E}[\Omega_i \hat{F}(H_i^*)] - \mathbb{E}[H_i^*]p_H^* - \left(\mathbb{E}[\Omega_i \hat{F}(\overline{H})] - \mathbb{E}[\overline{H}]p_H^*\right) \\ &= \mathrm{Cov}(\Omega_i, \hat{F}(H_i^*)) + \mathbb{E}[\Omega_i]\mathbb{E}[\hat{F}(H_i^*)] - \mathbb{E}[H_i^*]p_H^* - \left(\mathbb{E}[\Omega_i \hat{F}(\overline{H})] - \mathbb{E}[\overline{H}]p_H^*\right) \\ &= \mathrm{Cov}(\Omega_i, \hat{F}(H_i^*)) + \mathbb{E}[\Omega_i]\mathbb{E}[\hat{F}(H_i^*) - \hat{F}(\overline{H})] - p_H^* \left(\mathbb{E}[H_i^*] - \overline{H}\right), \end{aligned} \tag{1}$$

which is the expression in the text.

Now we prove the claim that there always exists a ration such that $S^P - S^R = \mathrm{Cov}(\Omega_i, \hat{F}(H_i^*))$. The second and third terms on the right hand side can be rearranged as

$$\mathbb{E}[\Omega_i]\mathbb{E}[\hat{F}(H_i^*)] - p_H^* \mathbb{E}[H_i^*] \tag{2}$$

$$+\mathbb{E}[\Omega_i]\mathbb{E}[\hat{F}(\overline{H})] - p_H^* \overline{H}, \tag{3}$$

where (2) gives the expected surplus across all farmers' input choices, when evaluated at the productivity of the mean farmer, and (3) gives the surplus under rationing at the productivity of the mean farmer. The surplus (2) is a constant that does not depend on $\overline{H}$ and may reasonably be assumed to be positive under the optimal Pigouvian regime. The surplus (3) is zero at $\overline{H} = 0$, initially increasing and concave. If we evaluate the difference between (2) and (3) at $\overline{H} = \mathbb{E}[H_i^*]$, we have

$$\mathbb{E}[\Omega_i]\mathbb{E}[\hat{F}(H_i^*)] - p_H^* \overline{H} - \left(\mathbb{E}[\Omega_i]\mathbb{E}[\hat{F}(\overline{H})] - p_H^* \overline{H}\right) = \mathbb{E}[\Omega_i]\left(\mathbb{E}[\hat{F}(H_i^*)] - \hat{F}(\mathbb{E}[H_i^*])\right) < 0,$$

where the inequality follows from the concavity of $F(\cdot)$. Therefore in the case where (2) is positive,

it will be initially greater than (3), but become lesser at some value of $\overline{H}$. Since the profit function is continuous the intermediate value function implies there exists an $\overline{H}$ such that these two terms are equal. In a case where (2) is negative, the Inada conditions on $\hat{F}(\cdot)$ again would imply the existence of such an $\overline{H}$, which must additionally be unique.

b Decomposition of error covariance into productivity and measurement error

Let inputs be divided into two sets, J for inputs taken as endogenous and J' for inputs taken as exogenous, such as inputs set by a binding ration. Under the Cobb-Douglas production function specified, expected log output can be written

$$z_{ic} = y_{Eic} \quad = \quad \frac{1}{1-\sum_{j\in J}\alpha_j}\left(\omega_{Eic} + \sum_{j\in J'}\alpha_j \ln j_{ic} + \sum_{j\in J}\alpha_j \ln\left(\frac{\alpha_j}{p_{jic}}\right)\right).$$

The observed output at harvest is

$$y_{ic} \quad = \quad z_{ic} + \epsilon_{Yic}.$$

The factor demand equations for observed inputs are

$$j^o_{ic} \quad = \quad z_{ic} + \ln\alpha_J - \ln p_{Jic} + \epsilon_{Jic}$$

Observed factor demands depend on expected output, itself a measure of total factor productivity, as well as output elasticities, farmer-plot specific prices, and measurement error in inputs.

Gollin and Udry (2020) introduce a decomposition to separate measurement error from other determinants of observed input demands and output. The key idea is the assumption that $p_{Jic} = p_{Ji}$ across all crops and plots for a farmer, where crops are indexed by c. Farmers may face farmer-specific prices, for example a high price of labor if they have a small family, or a high price of capital if they are credit constrained, but these farmer-specific prices are common across crops and plots. Under this assumption, it is possible to identify the variance of measurement error using variation within a farmer across crops.

Define $\tilde{j} = j^o_{ic} - \overline{j}_i$ as the deviation of input use from its mean for a given farmer, and let the tilde serve as an analogous difference operator for other variables. Since output elasticities and

prices are common across plots,

$$\begin{aligned}\tilde{y}_{ic} &= \tilde{z}_{ic} + \tilde{\epsilon}_{Yic} \\ \tilde{j}_{ic} &= \tilde{z}_{ic} + \tilde{\epsilon}_{jic}.\end{aligned}$$

With measurement error that is mean zero in logs and independent of productivity shocks, we can estimate the variance of z_{ic} as:

$$\begin{aligned}\hat{\sigma}^2_\omega &= \text{Cov}\left(\tilde{y}_{ic} - (W_{Hic} - \overline{W}_{Hic})\beta_H, \tilde{j}_{ic}\right) \\ &= \text{Cov}\left(\tilde{z}_{ic} + \tilde{\epsilon}_{Yic}, \tilde{z}_{ic} + \tilde{\epsilon}_{jic}\right) \to \text{Cov}\left(\tilde{z}_{ic}, \tilde{z}_{ic}\right).\end{aligned}$$

The economic idea is that if variance in output across plots is truly due to productivity shocks, and not to measurement error, then output and inputs should covary. If we observe a high variance of output across crops within a farmer, but no corresponding variance in inputs, then we should conclude that most of the variance in output is driven by measurement error.

We implement this estimator using the covariance across plots within a farmer of output and capital. With this estimate of the variance of z_{ic} we recover the variance of measurement error for output and inputs as $\hat{\sigma}^2_{\epsilon_j} = \text{Var}(\tilde{j}_{ic}) - \hat{\sigma}^2_\omega$ and use the estimated measurement error to deflate our estimates of TFP. Let $\widehat{TFP}_a$ be the raw residual from estimating (10). We calculate

$$\text{Var}\left(\widehat{TFP}_c\right) = (\widehat{TFP}_a - \widehat{\sigma}^2_{\epsilon_Y} - \sum_j \widehat{\sigma}^2_{\epsilon_j}\widehat{\alpha}_j)$$

and then form the deflated estimate of farmer-crop productivity $\widehat{TFP}_c$ as

$$\widehat{TFP}_c = \overline{\widehat{TFP}_a} + (\widehat{TFP}_a - \overline{\widehat{TFP}_a})\sqrt{\text{Var}\left(\widehat{TFP}_c\right) / \text{Var}\left(\widehat{TFP}_a\right)}.$$

This procedure implicitly assumes that the within-farmer across-crop variance of measurement error is the same as the across-farmer-crop variance in measurement error.

c Transfer rules for counterfactuals

Let $H_i^{Current} \leq \overline{H}$ be the usage of each farmer under the current uniform rationing regime and H_i^{Pigou} be usage under the Pigouvian regime with a uniform price. The state's present net revenue

from power supply is

$$R^{Current} = \sum_i H_i^{Current} P_i (p_E - c_E).$$

per day, where p_E is the present, low price of power. Net revenue is negative because the price of power is below the cost of supply. The state's net revenue under the Pigouvian pricing regime R^{Pigou} is calculated with the same formula, but at the higher price of $p_E^* > c_E$, and will therefore be strictly positive. The budget available for reallocation to farmers is the difference between state expenditures under rationing and under the Pigouvian regime

$$\Delta R = R^{Pigou} - R^{Current}.$$

There are N farmers on the grid. Under a flat transfer, each farmer receives a transfer of $T_i = T = \Delta R/N$. Under a land or pump-based transfer, each farmer receives a transfer that is proportional to their observed landholdings or pump capacity. For example, let there be N farmers on the grid, with each farmer i having land L_i. Total land under cultivation is $L = \sum_i L_i$. Each farmer receives a transfer $T_i = (L_i/L)\Delta R$.

C Appendix: Robustness checks and auxiliary estimates

This section considers the robustness of our estimates of the effect of well depth on farmer profits.

a Robustness to alternative candidate instrument sets

We define five different candidate instrument sets, described in Table C3. All of the instrument sets consist of geological data from the Bhuvan Bhujal project, described in Section 2 of the paper and Appendix A.

Table C3
Definition of candidate instrument sets

	Fractures (1)	Rocks (2)	Aquifers (3)	Main (4)	Large (5)
Fractures	*Yes*		*Yes*	*Yes*	*Yes*
Rock type		*Yes*	*Yes*	*Yes*	*Yes*
Rock share		*Yes*	*Yes*	*Yes*	*Yes*
Aquifer type			*Yes*	*Yes*	*Yes*
Fractures^2				*Yes*	*Yes*
Rock share^2				*Yes*	*Yes*
Fractures $\times$ Rock share				*Yes*	*Yes*
Fractures^2 $\times$ Rock share					*Yes*
Rock share^2 $\times$ Fractures					*Yes*
Rock share^2 $\times$ Fractures^2					*Yes*
Size of instrument set	3	130	153	419	1728

This table defines the instruments contained in our candidate geological instrument sets. Broadly, the geological variables consist of data on rock fractures, rock types, rock shares and aquifer types. Data on fractures consists of variables that capture the distance between a farmer's location and the nearest water conductive fracture, and the total length of such fractures in a radius of one and five kilometres around the farmer's location. Data on rock types consists of sixty-five dummy variables which indicate the type of rock at the farmer's precise location. Rock share variables capture the share of a given rock type in a five kilometre radius around the farmer's precise position. Aquifer types are dummy variables which indicate the presence of any of twenty types of aquifers that are located at the farmer's precise location. The instrument sets are composed of these basic variables and some of their interactions, as explicitly enumerated in this table. The instrument set labelled "Main" is the one used in our principal regression specification. The estimates generated from our "Large" instrument set are also included in Table 2 and Table D9 for reference. The other instrument sets are used in our robustness specifications in Table C5.

The different candidate instrument sets are comprised of functions of three different categories of geological variables: rock types, aquifer types and fractures. Rock type are variables for the type of rock underlying an area, such as basalt or gneiss. These variables are expected to predict groundwater levels because rocks have different porosity and therefore allow groundwater to penetrate to different depths. Aquifer types are a classification of what geological feature in an area bounds

groundwater flow. Fractures are variables indicating where there are significant underground faults in rock formations, also called lineaments. These fractures are referred to in hydrogeology as secondary porosity, and are secondary in the sense that they formed after a rock was initially deposited in an area, through seismic activity for example. In hard rock aquifers like Rajasthan's secondary porosity is an important determinant of groundwater flow.

The smallest instrument set, "Fractures," consists of only the fracture instruments (Table C3, column 1). The "Rocks" instrument set consists of rock type variables (column 2). The "Aquifers" set consists of aquifer variables in addition to the variables included in the other two candidates instrument sets (column 3). The "Main" instrument set, which we use for our main estimates of the effect of depth on profits, includes the variables in the "Aquifers" set and the non-trvial first-order interactions of rock and fracture variables (column 4). The "Large" instrument set consists of the variables in the "Aquifers" set and the non-trivial second order interactions between fracture and rock variables (column 5). The candidate instrument sets vary in size, with the smallest set consisting of 3 variables and the largest containing 1728 variables.

Table C4 shows the results of the first-stage regression equations with each candidate instrument set. The top of the table shows what groups of instruments have any variables that are actually selected by LASSO. The bottom of the table shows summary statistics on the strength of the instruments and goodness of fit. All of the instrument sets have first-stage F-statistics of 30 or more, above typical critical values for weak instruments. (We report F-statistics to follow convention, but this statistic should be interpreted with caution, since it will no longer strictly follow an F-distribution when calculated post-variable selection (Lockhart et al., 2014)). The IV-PDS specifications and instrument sets achieve a lower prediction error than two-stage least squares with only three pre-selected instruments. As the set of candidate instruments passed to LASSO increases by a factor of more than ten, the number of instruments selected barely grows, supporting the sparsity assumption behind the IV-PDS estimator. In agreement with the geological literature, the LASSO procedure consistently selects a similar set of rock type variables and rock types interacted with fractures.

Tables C5 compares the instrumental variables estimates of our profit regressions obtained with different candidate instrument sets. In Table C5, Panel A, with total profit as the dependent variable, the estimated coefficient on well depth (measured in units of one standard deviation =

Table C4
First Stage: Well depth on instruments

	IV-2SLS	IV-PDS	IV-PDS	IV-PDS	IV-PDS
	Fractures	Rock	Aquifers	Main	Large
	(1)	(2)	(3)	(4)	(5)
Fractures	*Yes*				
Rock shares		*Yes*	*Yes*	*Yes*	*Yes*
Rock types		*Yes*	*Yes*	*Yes*	*Yes*
Aquifer types			*Yes*	*Yes*	*Yes*
Fractures^2					
Rock shares^2				*Yes*	
Fractures × Rock shares				*Yes*	*Yes*
Fractures^2 × Rock shares					*Yes*
Fractures × Rock shares^2					
Fractures^2 × Rock shares^2					*Yes*
RMSE	148.4	417.8	427.2	423.8	364.5
F	137.3	104.5	98.3	86.8	84.6
Candidate Instruments		130	153	419	1728
Instruments Selected		11	12	16	18
Unique Farmers	3999	3999	3999	3999	3999
Farmer-Crops	9540	8973	8973	8973	8973

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

187 feet) ranges from INR −7.01 thousand per Ha (standard error INR 2.70 thousand per Ha) to INR −8.87 thousand per Ha (standard error INR 2.47 thousand per Ha) in IV-PDS specifications, regardless of whether the instrument set includes only rock types, only aquifer types, the main instrument set or the full instrument set. Depending on the size of the candidate instrument set, between 11 and 19 instruments are selected by LASSO to have non-zero coefficients. With cash profit as the outcome variable, in panel B, the coefficient on depth is consistently larger, but also relatively stable across specifications. For either profit outcome, fixing a small instrument set, based only on fractures, and estimating the profit equation via two stage least squares yields very imprecise estimates (column 1), showing the value of the IV-PDS method for improving precision in our application.

We conclude that: (a) geological factors have a strong first stage for the prediction of well depth; (b) the LASSO procedure selects similar instruments from widely varying sets of candidate instruments; (c) the number of instruments selected does not grow with size, supporting the sparsity

Table C5
Hedonic regressions of profit on well depth (robustness to different instrument sets)

	IV-2SLS	IV-PDS	IV-PDS	IV-PDS	IV-PDS
	Fractures	Rock	Aquifers	Main	Large
	(1)	(2)	(3)	(4)	(5)
Panel A. Total Profit, reported ('000 INR per Ha)					
Well depth (1 sd = 187 feet)	−1.06	−7.95***	−8.39***	−8.44***	−6.95***
	(17.5)	(2.58)	(2.59)	(2.41)	(2.63)
Mean dep. var	−5.06	−5.06	−5.06	−5.06	−5.06
Candidate Instruments	3	130	153	419	1728
Instruments Selected		11	12	16	18
Unique Farmers	3999	3999	3999	3999	3999
Farmer-Crops	8973	8973	8973	8973	8973
Panel B. Cash Profit ('000 INR per Ha)					
Well depth (1 sd = 187 feet)	−45.8*	−12.2**	−16.2***	−10.7**	−18.1***
	(25.8)	(5.70)	(5.98)	(5.03)	(6.06)
Mean dep. var	−13.5	−13.5	−13.5	−13.5	−13.5
Candidate Instruments	3	130	153	419	1728
Instruments Selected		4	5	9	6
Unique Farmers	2121	2121	2121	2121	2121
Farmer-Crops	3243	3243	3243	3243	3243

This table shows instrumental variable regressions of different measures of agricultural output on farmer well depth. The data is from the main agricultural household survey and the observations are at the farmer-by-crop level. The dependent variable changes by panel. In Panel A, the dependent variable is reported total profit (INR per Ha), in Panel B, it is cash profit. All model specifications control for the toposequence (elevation and slope), along with subdivisional and plot size effects, as defined in Table 2. The set of candidate instruments changes by column; the definitions of different instrument sets used in the model specifications above can be found in Table C3. Standard errors are clustered at the feeder, the primary sampling unit. The statistical significance of a coefficient at certain thresholds is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

assumption of (Belloni et al., 2012); (d) our findings on the effect of well depth on profits are robust to the use of different candidate instrument sets.

b Robustness to inclusion of controls

The exclusion restriction is that, conditional on included exogenous controls, the geological variables used as instruments do not have a direct effect on farmer profits, other than through their effect on groundwater levels. Here we consider how our instrumental variables estimates vary depending

on the set of control variables for surface productivity included in the structural equation.

We consider five different types of controls in our analysis: subdivisional effects, plot size decile effects, toposequence, soil quality controls, and temperature controls. Subdivisional effects are dummy variables for each of the six subdivisional-office areas from which farmers were sampled. Plot size decile effects are dummy variables which indicate the decile of the plot size distribution within which a particular farmer falls. We include weather controls in some specifications, though in our data, which covers a single season, there is very little measured variation in weather, especially after we condition on subdivision fixed effects. All of these farmers face similarly hot conditions with negligible rainfall during the Rabi season (See Appendix A e).

Table C6 holds constant the IV-PDS estimation method and candidate instrument set and varies the set of exogenous controls included in the specification. Column 1 includes only SDO fixed effects, column 2 adds plot size effects, column 3 adds toposequence, column 4 adds soil quality controls and column 5 adds temperature controls. Panel A reports outcomes for total profit and panel B for cash profits. In both panels, the coefficients vary little across specifications, and are generally within one standard error of our main estimate and typically even closer. At the same time, the controls themselves have significant effects on profits, for example, profits are lower in steeper areas and the soil quality controls are jointly significant (not reported). We conclude that the instrumental variables based on underground geology are not highly correlated with observable determinants of productivity on the surface.

c First-stage estimates for production function

Table 3 in the main text reports estimates of the production function. Table C7 reports estimates of the first-stage equations for the instrumental variables estimates in Table 3, column 2. Each column of the table has as the dependent variable the logarithm of one farmer-crop input and the independent variables the superset of all instruments. We suppress reporting the coefficients on the geological instruments for brevity.

The instruments are described in the Table 3 notes in brief. Section 4 A describes the geological instruments. Section 5 A describes the other instruments. Table C8 provides summary statistics on the non-geological instruments. These include the size of farmer parcels; the number of adult males in the household; and the median seed price in a household's feeder.

Table C6
Hedonic regressions of profit on well depth (robustness to inclusion of controls)

	(1) IV-PDS	(2) IV-PDS	(3) IV-PDS	(4) IV-PDS	(5) IV-PDS
Panel A. Total Profit ('000 INR per Ha)					
Well depth (1 sd = 187 feet)	-9.66^{***}	-8.91^{***}	-8.83^{***}	-8.44^{***}	-6.03^{**}
	(2.61)	(2.75)	(2.63)	(2.41)	(2.57)
Subdivisional effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Plot size effects		*Yes*	*Yes*	*Yes*	*Yes*
Toposequence			*Yes*	*Yes*	*Yes*
Soil quality controls				*Yes*	*Yes*
Temperature					*Yes*
Mean dep. var	−5.06	−5.06	−5.06	−5.06	−5.06
Candidate Instruments	419	419	419	419	419
Instruments Selected	16	15	15	16	14
Unique Farmers	4008	4008	3999	3999	3999
Farmer-Crops	8991	8991	8973	8973	8973
Panel B. Cash Profit ('000 INR per Ha)					
Well depth (1 sd = 187 feet)	-11.5^{**}	-9.91^{*}	-9.28^{*}	-10.7^{**}	-10.5^{**}
	(5.06)	(5.08)	(4.77)	(5.03)	(4.69)
Subdivisional effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Plot size effects		*Yes*	*Yes*	*Yes*	*Yes*
Toposequence			*Yes*	*Yes*	*Yes*
Soil quality controls				*Yes*	*Yes*
Temperature					*Yes*
Mean dep. var	−13.5	−13.5	−13.5	−13.5	−13.5
Candidate Instruments	419	419	419	419	419
Instruments Selected	10	9	12	9	8
Unique Farmers	2127	2127	2121	2121	2121
Farmer-Crops	3253	3253	3243	3243	3243

This table shows instrumental variable regressions of different measures of agricultural output on farmer well depth. The data is from the main agricultural household survey and the observations are at the farmer-by-crop level. The dependent variable changes by panel. In Panel A, the dependent variable is reported cash profit (INR per Ha), in Panel B, it is total profit which is inclusive of the value of the farmer's own consumption (INR per Ha). All models use the main instrument set as described in Table C3. The set of controls included changes by column; for example, the first column only includes subdivisional effects whereas the last column includes all five sets of controls considered. Standard errors are clustered at the feeder, the primary sampling unit. The statistical significance of a coefficient at certain thresholds is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C7
First stage estimates from production function estimation

	log(Water) (1)	log(Labor) (2)	log(Land) (3)	log(Capital) (4)
Size of the largest parcel (Ha)	0.19***	0.21***	0.30***	0.24***
	(0.011)	(0.0083)	(0.0082)	(0.0081)
Size of the 2nd largest parcel (Ha)	1.17***	0.99***	1.39***	1.09***
	(0.057)	(0.041)	(0.041)	(0.040)
Size of the 3rd largest parcel (Ha)	0.75***	0.78***	1.07***	0.93***
	(0.13)	(0.094)	(0.093)	(0.092)
Size of the largest parcel squared (Ha^2)	−0.0032***	−0.0045***	−0.0067***	−0.0053***
	(0.00038)	(0.00028)	(0.00028)	(0.00027)
Size of the 2nd largest parcel squared (Ha^2)	−0.28***	−0.23***	−0.32***	−0.24***
	(0.021)	(0.015)	(0.015)	(0.015)
Size of the 3rd largest parcel squared (Ha^2)	−0.063	−0.18***	−0.26***	−0.23***
	(0.058)	(0.042)	(0.042)	(0.041)
Adult males	0.048***	0.070***	0.032***	0.024***
	(0.012)	(0.0084)	(0.0083)	(0.0082)
Adult males squared	−0.00100	−0.0025***	−0.00094	−0.00068
	(0.00081)	(0.00059)	(0.00058)	(0.00058)
Seed price ('00 INR/kg)	−0.12***	−0.13***	0.045***	−0.070***
	(0.022)	(0.016)	(0.016)	(0.016)
Seed price squared ('0,000 INR^2/kg^2)	0.019***	0.016***	−0.0047*	0.0048*
	(0.0038)	(0.0028)	(0.0027)	(0.0027)
Geological variables	*Yes*	*Yes*	*Yes*	*Yes*
Mean dep. var	6.70	3.66	−0.78	2.53
R^2	0.24	0.31	0.44	0.35
F-statistic	101.4	155.4	283.7	182.9
Farmers	3998	3998	3998	3998
Farmer-crops	8711	8711	8711	8711

This table reports coefficients of the first stage equation for each input in the the instrumental variables estimates of the production function regression. Each column has as the dependent variable the logarithm of farmer-crop inputs and the independent variables the superset of all instruments. There are four sets of instruments. (i) The size of the farmer's three largest parcels owned and size squared. (ii) The number of adult males in the household and the number of adult males squared. (iii) The mean price of seeds in the farmer's feeder and the mean price squared, where each variable leaves out the farmer's own prices paid. (iv) Geological variables that influence groundwater depth. All specifications include controls for toposequence (slope and elevation), subdivisional fixed effects and village-level soil quality indicators. Standard errors are clustered at the feeder, the primary sampling unit. Statistical significance at certain thresholds is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C8
Summary statistics for non-geological instruments

	Mean (1)	Std. dev (2)	25th (3)	Median (4)	75th (5)	Farmer-crops (6)
Land instruments						
Size of the largest parcel (Ha)	1.14	1.73	0.38	0.73	1.36	8711
Size of the 2nd largest parcel (Ha)	0.14	0.38	0	0	0.081	8711
Size of the 3rd largest parcel (Ha)	0.030	0.16	0	0	0	8711
Labor instruments						
Adult males	1.66	1.42	1	1	2	8711
Capital instruments						
Seed price ('00 INR/kg)	1.64	1.97	0.31	0.58	2.45	8711

This table provides summary statistics on the instruments used to generate exogenous variation in productive inputs for production function estimation. All observations are at the farmer-crop level. The first block of summarize land instruments, which consist of the size of the three largest plots of land owned by a farmer. The second block summarizes the main instrument for labor, which is the number of adult males in the household. Finally, the last block summarizes seed prices which affect capital inputs exogenously, assuming the farmet has limited market power. Seed prices for each farmer-crop observation is calculated as the median price of all seed inputs in the feeder in which the farmer is located. Geological instruments are excluded from this summary since they are numerous and heterogenous, and their units are not always easy to interpret.

D Appendix: Additional Results

a Regulation binds on intensive margin

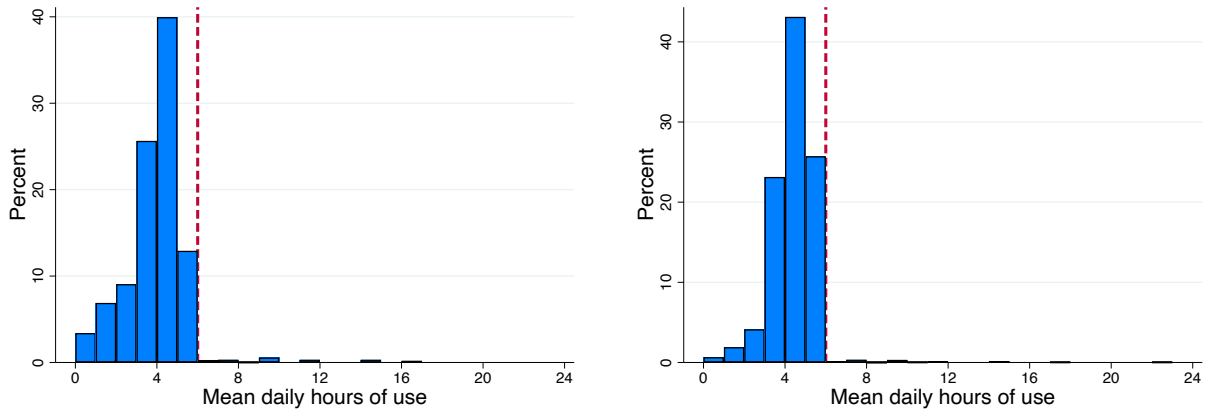
Section 2 D argues that the distributions of hours of power supply and hours of power use, shown in Figure 2, are evidence that the ration of power supply is binding, despite that there is a small gap between average hours of supply and hours of use. Here we provide additional evidence for two reasons that may account for this gap: farmers needing to turn on their pumps, and farmers reporting average values of use over the whole season, including some periods when irrigation is less intensive.

The first reason to expect a gap between supply and use is that farmers need to turn on their pumps when the grid begins to supply power. The hours of supply are not always predictable, and farmers may live at some distance from their fields. In our data 49% of farmers have auto-starters that turn on their pumps automatically when the power comes on. Figure D5 shows the distribution of hours of use for farmers without (panel A) and with (panel B) auto-starters. The share of farmers reporting average use of at least five hours per day over the whole season is 12 percentage points higher for farmers with autostarters. Therefore there is evidence that the gap between hours of use and hours of supply is due in part to farmers needing to turn on their pumps manually.

The second, more important, reason why hours of use on average over the dry season is below hours of supply is that farmers do not irrigate continuously for the whole season. In particular, there is a period of peak plant growth, when dry season crops must be irrigated intensively, but also two shoulder periods after planting and before harvest when less water is needed. Our survey data applies to the entire season, but it is possible to see this peak in aggregate data sources with monthly observations over the course of the same season. Figure D6 provides evidence that these shoulder periods should be expected to reduce average water demand over the whole dry season. The blue bars show agronomic estimates of the monthly water demand for the dry season wheat crop in Haryana, adjacent to Rajasthan (Pakhale, Prasun and Nale, 2010). The solid line shows the monthly average energy injection in agricultural electricity feeders in our study area, from administrative data on grid power supply obtained from the electric utility.

The main finding from the figure is that peak demand in the dry season consists of three truly

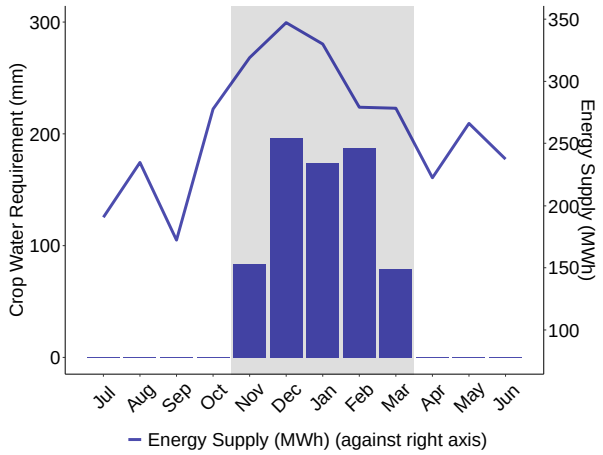
Figure D5: Hours of electricity use by auto-starter status



A. Farmers without autostarters B. Farmers with autostarters

The figure shows power rationing data from our Rajasthan survey, as in Figure 2 but broken down by whether farmers have autostarters or not. Panel A shows the distribution of the average hours of supply per day during the Rabi season of 2016-2017 for farmers with autostarters installed. Panel B shows the same distribution for farmers without autostarters.

Figure D6: Seasonality of peak energy and water demand



The figure shows data on the pattern of peak consumption of energy and water during the Rabi (dry) season in Rajasthan. The blue bars show agronomic estimates of the monthly water demand for the dry season wheat crop in Haryana, adjacent to Rajasthan (Pakhale, Prasun and Nale, 2010). The solid line shows the monthly average energy injection in agricultural electricity feeders in our study area, from administrative data on grid power supply obtained from the electric utility.

peak months, from December to February, and two shoulder months of November and March. Both the agronomic predicted water demand and the actual energy injection in agricultural feeders follow the same pattern. The peak energy demand (or water demand) is sustained over the period from December through February but the shoulder months of November and March have elevated, but slightly lower, levels of demand. This seasonal pattern of usage likely accounts for farmers

reporting less than six hours of use over the full season despite being constrained at the true peak. For example, if we take true peak during December to February as representing six hours of pump use, then the water demand series implies that the corresponding average use over the full season, including the shoulder months, would be 4.7 hours, consistent with the modal value of hours of use in Figure 2, panel B.

b Regulation binds on extensive margin

Farmers may in principle evade the rationing of power supply by connecting more or larger pumps, to extract more water during the ration of 6 hours. Or, farmers could run single-phase pumps, to extract water during the 18 hours when power supply has only one phase. The state utility regulates the type, number and capacity of pumps on agricultural electricity connections to prevent such evasion. This subsection presents evidence that these regulations also bind, so that the rationing regime as a whole does act as a limit on water use.

Consider first the margin of the type of pump. The ration is enforced by switching off two phases of three-phase alternating current electricity supply (see footnote 9). In principle, one could run a single-phase pump, the kind used in some small household appliances, outside of the rationed hours to continue extracting water at a reduced rate. The distribution company disallows this by specifying in the tariff order that all agricultural electricity connections are for "AC 50 Hertz, Three Phase 400 V" pumpsets. The Superintending Engineer (Commercial) for JVVNL confirmed to us that JVVNL will not provide agricultural electricity connections to farmers to run single-phase pumps.

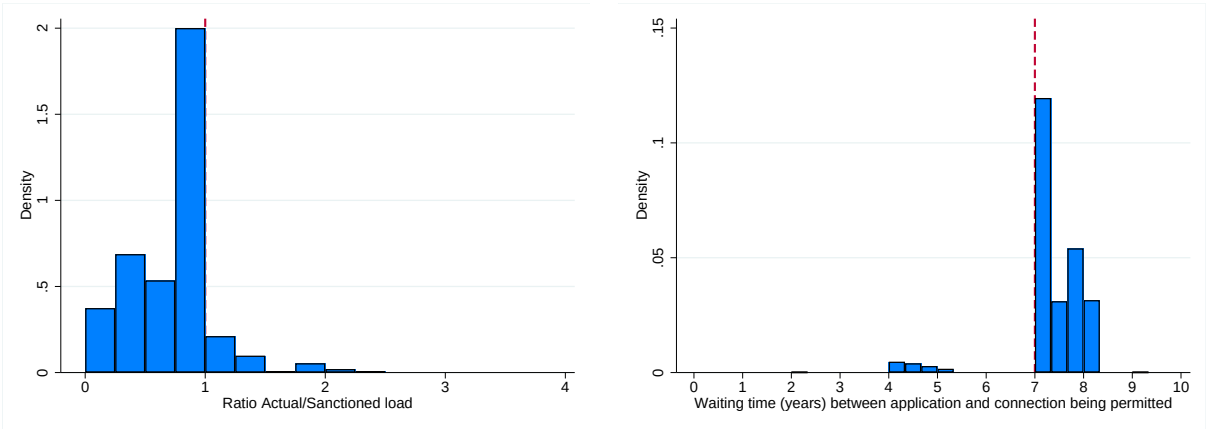
Consider next the margin of farmers adding more pumps. To get another pump, farmers have to apply to get a new agricultural connection. The number of connections is limited by rationing the number of applications that are granted off of the waiting list. We collected administrative data on the waiting list including the time of initial application and the time that applications were granted. Figure D7, panel A shows the distribution of the gap between the applications and their clearance. At the time of our data collection, the waiting list was long enough that farmers who had applied 7 or 8 years prior were just getting connections approved, and very few farmers who applied later had their connections approved. This waiting list mechanism therefore serves as a ration on the extensive margin of number of pumps connected to the grid.

Consider finally the margin of farmers adding larger pumps. When a farmer is given an agricultural electricity connection, that connection specifies a "sanctioned load," that is, the size of pump that a farmer is permitted to run on that connection. The sanctioned load may differ depending on areas and land size. Our sampling frame contains data on sanctioned load and our surveyed asked farmers about actual load, so that we can compare the two to look for evidence of evasion. Figure D7, panel B shows that most farmers use exactly their sanctioned load, or in some cases have smaller pumps, but seldom larger ones. This suggests that the sanctioned load regulation is enforced. We conclude that farmers cannot evade the ration on the number of hours of supply by connecting more pumps or a greater number of pumps.

Figure D7: Extensive margin

A. Distribution of waiting times to acquire pump connection

B. Ratio of actual pump load to sanctioned load



This figure provides empirical evidence that the ration binds on all dimensions. Panel A shows the distribution of the ratio of the actual pump load in our farmer survey to the sanctioned load, which is the load the farmer is allowed by the government to have by the terms of their electricity connection. The modal farmer reports that they use exactly the sanctioned load and relatively few farmers have actual pump loads above the sanctioned load. Panel B shows the distribution of wait times in years for acquiring an agricultural pump connection from the power utility company in Hindoli and Mundawar, two of the subdivisional areas in our sample. The data consists of application and approval dates of connection requests from farmers who applied for a pump between 2010 and 2014.

c Adaptation to environmental change

The results in the main text show that deeper wells decreases farmer profits. Farmer responses to groundwater scarcity may be complex. This subsection presents results for additional outcome variables to characterize why profits decline.

We estimate that profits fall in part because farmers with deeper wells produce less output.

Table D9 presents results for yield (panel A) and the total value of output (panel B). Yield is measured in quintals (100 kg units) per Ha and aggregated across crops, regardless of their value. The panel A, column 3 estimate is that farmer yields decline by 0.054 quintals per Ha (standard error 0.012 quintals per Ha), where the mean of the dependent variable is 46.3 quintals. Thus a one standard deviation increase in well depth would decrease yield by 10 quintals per Ha, about 20% of the mean yield. The corresponding result from panel B is an INR 48.0 per Ha (standard error INR 11.4 per Ha) decrease in the value of output, or 14% of the mean value of output per standard deviation increase in depth.

Farmers use a range of irrigation technologies and techniques in order to deliver the water they extract from the ground to their crops. We next examine whether changing irrigation techniques can compensate for groundwater scarcity. Table D10 uses the same identification strategy developed to estimate the effect of water scarcity on profits to study how farmers endogenously respond to a lack of water. We consider responses on a number of margins that are likely related to the intensity and efficiency of water use: whether a farmer plants a high-yielding variety of crop, which requires more water; whether a farmer levels his parcels before planting, which conserves water; whether a farmer uses sprinkler irrigation, which conserves water; whether a farmer instead uses furrow or flood irrigation, which is a relatively wasteful technique; and whether a farmer reports the crop on a given plot was under-irrigated.

The main finding of Table D10 is that farmers adapt to water scarcity by disinvestment in both water intensity and in water efficient methods. On average 62% of farmers plant a high-yielding variety of crop. Increasing water depth by one standard deviation (187 feet) reduces the probability of planting a high-yielding variety of crop by 9% (column 1, -0.049 / 0.62, standard error 5.4 pp). The same decline in water reduces the probability a parcel is leveled by 8 percentage points (standard error 4.4 pp), or 39% (column 2). It reduces the probability of using sprinkler irrigation by 10 percentage points (standard error 4.1 pp), or 33% (column 3) and appears to increase the probability of furrow or flood irrigation, an alternative technique that uses more water. Finally, it sharply increases the probability that a farmer reports their crop was under-irrigated, by 12 percentage points, or 62% on a base of 19 percentage points.

We interpret this consistent pattern as showing that farmers do adapt to water scarcity, but adapt by disinvestment rather than investment. This adaptation can be rationalized if the avail-

ability of water is complementary to water saving techniques. For example, suppose that sprinkler irrigation technology has some fixed cost but acts literally as a factor multiplier on water, such that the amount of water delivered to crops is αW for water extraction W, and $\alpha_{Sprinkler} > \alpha_{Furrow}$. Then farmers may wish to invest in water saving only if there is enough water to be worth saving. The Green Revolution intensified the input bundle that farmers used to include more capital, more intermediates like fertilizer and more water. A scarcity of water, in our estimates, reverses this intensification.

d Marginal social benefit and cost of an increased ration

Figure 4 compares the marginal benefit and marginal cost of an increase in the ration using our estimates of the effect of depth on profits. This subsection gives the calculations underlying the results in this figure. Equation (6) gives the marginal benefit and marginal cost of increasing the ration. We decompose the marginal benefit using the estimated effect of depth on profits as shown in equation (2).

Table D11, column 1, panel A carries forward our preferred estimate of a INR 8.87 thousand per Ha decrease in profit per standard deviation of depth (Table 2, Panel A, column 3). The estimated *decrease* of profits with depth—deeper water lowers water input, for a fixed ration—is equivalent to an *increase* in profits of INR 2200 per Ha for one additional hour of power supply (standard error INR 623 per Ha per hour) (Table D11, panel A, column 1).[1] The marginal private cost of increasing the ration, which is the cost only of the additional power that farmers would consume, is estimated to be INR 1300 per Ha-hour (Table D11, column 2, panel A). The marginal social cost of INR 2300 per Ha-hour additionally includes the opportunity cost of water (equation 2; reported in column 2, panel B).

[1] This estimate applies the average value $\overline{D}/\overline{H}$ to an equally-weighted regression. We have also estimated a version of (7) weighted by H_i/D_i, to be strictly consistent with (6), and find extremely similar results.

Table D9
Hedonic regressions of yield on well depth

	OLS (1)	OLS (2)	IV-PDS (3)	IV-PDS (4)
Panel A. Yield (quintals per Ha)				
Well depth (1 sd = 187 feet)	−7.99***	−2.20	−7.89***	−4.49
	(1.12)	(1.40)	(2.84)	(2.87)
Toposequence		*Yes*	*Yes*	*Yes*
Soil quality controls		*Yes*	*Yes*	*Yes*
Subdivisional effects		*Yes*	*Yes*	*Yes*
Plot size effects		*Yes*	*Yes*	*Yes*
Mean dep. var	46.3	46.3	46.3	46.3
Candidate Instruments			419	1728
Instruments Selected			14	17
Unique Farmers	4013	4004	4004	4004
Farmer-Crops	9554	9536	9536	9536
Panel B. Total Value of Output, imputed ('000 INR per Ha)				
Well depth (1 sd = 187 feet)	−0.51	−2.95**	−8.64***	−6.38**
	(1.05)	(1.33)	(2.62)	(2.87)
Toposequence		*Yes*	*Yes*	*Yes*
Soil quality controls		*Yes*	*Yes*	*Yes*
Subdivisional effects		*Yes*	*Yes*	*Yes*
Plot size effects		*Yes*	*Yes*	*Yes*
Mean dep. var	65.4	65.4	65.4	65.4
Candidate Instruments			419	1728
Instruments Selected			14	17
Unique Farmers	4009	4000	4000	4000
Farmer-Crops	9290	9272	9272	9272

The table reports coefficients from regressions of agricultural output measures on well depth and controls. The data is from the main agricultural household survey and the observations are at the farmer-by-crop level. The dependent variable changes in each panel. In Panel A, the dependent variable is yield (quintals per Ha). In Panel B, the dependent variable is the value of output (INR per Ha), where the price for each crop is taken to be the median of the price reported at the SDO level. Well depth is the reported depth of a given farmer's well. Toposequence includes controls for elevation and slope. Subdivisional effects are dummy variables for each of the six sub-divisional offices of the distribution company from which farmers were sampled. Plot size effects are dummy variables indicating the plot size decile for each farmer-crop based on its plot area. Standard errors are clustered at the feeder level, the block within which farmers were randomly selected. The statistical significance of a coefficient at certain thresholds is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D10
Instrumental variable estimates of farmer adaptation to water scarcity

	IV-PDS	IV-PDS	IV-PDS	IV-PDS	IV-PDS
	High-yielding variety	Parcel leveled	Sprinkler irrigated	Furrow/Flood irrigated	Under irrigated
	(1)	(2)	(3)	(4)	(5)
Well depth (1 sd = 187 feet)	−0.049	−0.080*	−0.099**	0.051	0.12***
	(0.034)	(0.044)	(0.041)	(0.042)	(0.031)
Mean dep. var	0.62	0.19	0.30	0.34	0.20
Candidate Instruments	419	419	419	419	419
Instruments Selected	11	10	11	11	10
Unique Farmers	3998	3982	4006	4006	3982
Farmer-Crops	8711	6857	9748	9748	6857

This table shows instrumental variable regressions of potential margins of adaptation to water scarcity on farmer well depth. Each column presents estimates from a model with a different outcome variable, as shown in the column headers. The data is from the main agricultural household survey and the observations are at the farmer-by-crop level for all but the first column where the data is at the farmer-by-parcel level. All the model specifications control for the toposequence (elevation and slope), along with subdivisional and plot size effects, as defined in Table 2. We use our preferred candidate instrument set which is labelled Main in Table C3 . Standard errors are clustered at the feeder, the primary sampling unit. The statistical significance of a coefficient at certain thresholds is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D11
Optimality of Ration

	Marginal benefit (1)			Marginal cost (2)	
				Private cost	
$-d\Pi/dD$	8.440	INR 000s per Ha-sd	$dE/d\overline{H}$	246	kWh per Ha-hr
$\times\overline{D}/\overline{H}$	0.25	sd / hr	$\times(c_E - p_E)$	5.30	INR per kWh
$d\Pi/d\overline{H}$	2.091	INR 000s per Ha-hr	$dPC/d\overline{H}$	1.304	INR 000s per Ha-hr
				Opportunity cost	
			$dW/d\overline{H}$	0.39	liter 000s per Ha-hr
			$\times\lambda_w$	3.248	INR per liter 000s
			$dOC/d\overline{H}$	1.254	INR 000s per Ha-hr
				Social cost	
			Private	1.304	INR 000s per Ha-hr
			+Opportunity	1.254	INR 000s per Ha-hr
			Social	2.558	INR 000s per Ha-hr

The table compares the marginal benefit and marginal cost associated with a one hour increase in the ration of electricity. Column 1 gives the marginal benefit of the increase in the ration calculated using equation (6). We identify the average efffect of water depth on profits using the specification shown in column 3 of Table 2. We weight by the ratio of the averages $\overline{D}_i/\overline{H}_i = 46.2$ since it is essentially the same as the average of the ratios $\overline{D_i/H_i} = 46.9$. Column 2 gives the marginal cost of the increase in the ration. The private marginal cost is the marginal cost of generating and distributing power. The opportunity cost is the external cost of water extraction. The social cost is the sum of the private marginal cost and the opportunity cost of water. See the right-hand side of equation 2 for the expression. We deduct here the small price of electricity that farmers already pay, since this small price is accounted for in farmer profits.

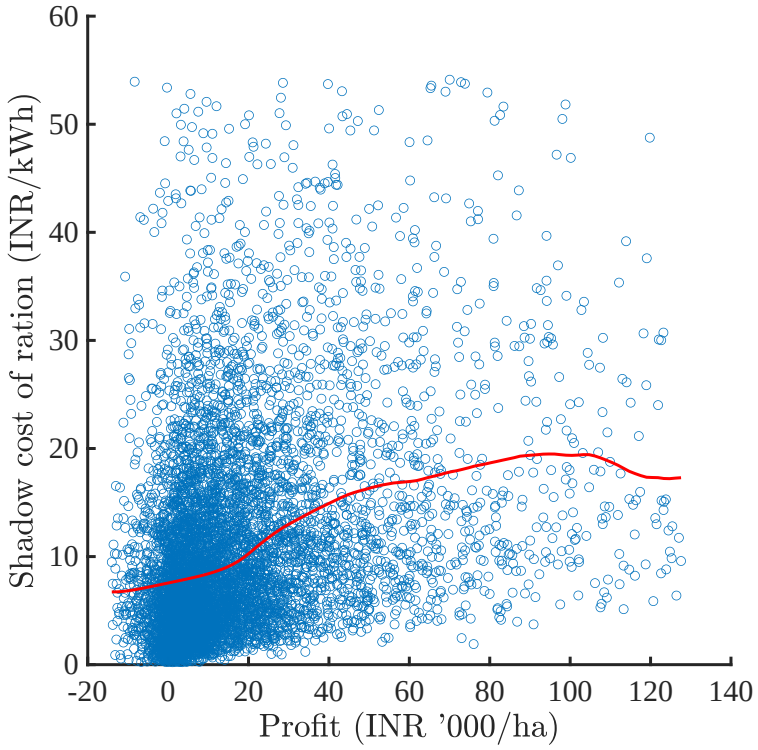
e Additional results from the production model and counterfactuals

Section 5 argued that farmers with higher ex ante profits were more likely to benefit from Pigouvian reform, before transfers, because more productive farmers had a higher shadow cost of the ration. This part provides more direct evidence on this point from the shadow cost of the ration and the patterns of input usage in counterfactual simulations.

Figure D8 plots the shadow cost of the ration against the ex ante profitability of a farmer in the rationing regime. Every point represents a farmer-crop observation. The shadow cost, the distribution of which is plotted in the paper (Figure 6), is the price of electricity such that a farmer would optimally choose to use the amount of power they receive under the ration. The shadow cost depends on the farmer's productivity, well depth, pump size and factor endowments. Figure D8 shows that more profitable farmers have much higher shadow costs of the ration. The average shadow cost more than doubles between the least and most profitable farmers. For example, a farmer with profits near zero would be predicted to have an average shadow cost around INR 8 per kWh, whereas a farmer with profits of INR 60 thousand per Ha would have an average shadow cost around INR 18 per kWh. At any given level of profitability, there is nonetheless large variability in the shadow cost.

The pattern of reform aiding larger farmers is also visible in the predicted counterfactual patterns of input choices under the Pigouvian regime, as compared to modeled input choices in the rationing regime. Figure D9 plots log input usage and log output for each farmer-crop under the Pigouvian regime against the same variable under the rationing regime. We include the endogenous inputs of hours of electricity use (panel A), water use (panel B) and capital use (panel C), as well as output (panel D), and omit the exogenous inputs of land and labor since they are held fixed in our main counterfactual simulations. (The maximal water use, under the Pigouvian regime, is truncated for farmers who would be predicted to run their pumps all 24 hours in a day.) The patterns of input use and output show a strong, positive correlation between input use across regimes, which is sensible, because farmers with higher factor endowments of land and capital, higher productivity and shallower wells will be expected to use more inputs and produce more outputs in either regime. Additionally, the data show an upward rotation of the pattern of input use and output. For example, consider the pattern of capital use (panel C). For farmers with initially low

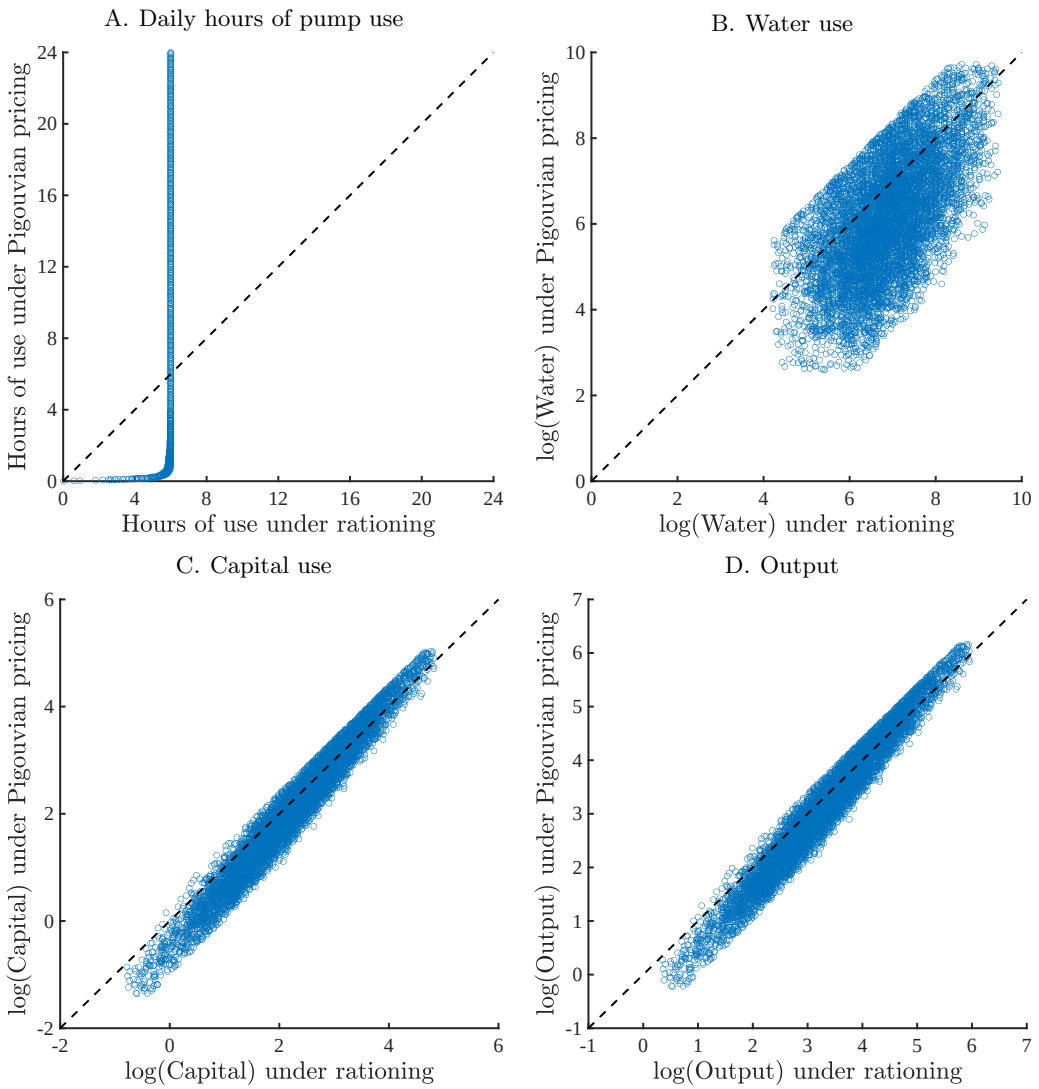
Figure D8: Scatterplot of the shadow value of the ration against profit



The figure shows a scatterplot of the relationship between farmer profits and the shadow cost of the ration, calculated as the price at which an unconstrained farmer would choose to use the rationed amount of power. The data is at the farmer-crop level, averaged over 200 counterfactual simulations with different productivity draws for each farmer-crop observation in the sample. The horizontal axis shows the average total profit for a farmer-crop observation across all simulations, denoted in thousands of rupees per hectare. The vertical axis shows the corresponding shadow cost of ration, in rupees per kilowatt hour of energy. The red line shows a kernel regression estimate of the shadow value as a function of profits.

levels of capital under rationing, capital use falls below the 45-degree line under the Pigouvian regime. For farmers with higher levels of capital, capital use is predicted to increase under the Pigouvian regime. The pattern for output is the same (panel D). These plots therefore illustrate the mechanism by which initially productive farmers or farmers with high factor endowments see greater increases in production when the ration is replaced with Pigouvian pricing.

Figure D9: Inputs and outputs under the Pigouvian and rationing regimes



This figure shows input choices and farm output under the status quo against inputs and outputs under a Pigouvian pricing regime. The data is at the farmer-crop level, aggregated from the counterfactual simulations consisting of 200 productivity draws for each farmer-crop observation in the sample. Panel A shows the daily hours of pump use in rationing and Pigouvian regimes. Panel B shows the logarithm of water used (in thousands of liters) in the two regimes. Panel C shows the logarithm of the total value of capital used as production inputs. Finally, panel D shows the logarithm of the total value of output in the Rabi season.

E Appendix: Opportunity Cost of Water (Not for Publication)

The model we present in the main text is static, but optimal groundwater policy is a dynamic problem (Timmins, 2002). Water extracted today lowers the groundwater level tomorrow, which increases the cost of extraction in the future or lowers the amount of water extracted, for fixed extraction effort. The cost of water extraction today is therefore a pure opportunity cost, which can be measured by the effect of today's extraction on the present discounted value of future profits.

In this section we therefore present a simplified, dynamic version of our main model in order to calculate the opportunity cost of water. This model has two parts. First, the production function, for which we use the parameters of our estimated production function applied to a single, representative farmer with average levels of productivity and input usage. Second, a law of motion for how water use affects groundwater depletion and therefore future water depths.

The model is dynamic because the state variable, water depth, connects present water extraction to future extraction and profits. The representative farmer in the model, however, does not make forward-looking decisions about present extraction. That is because the purpose of this model is not to represent farmer input and production decisions, at the aggregate level, but rather to value the effect of current water use on future profits. We represent farmer input and production decisions in our main model, in Section 3, at the farmer-plot level. At that level of disaggregation it is reasonable to assume that farmers do not internalize the external costs of their own water use.

a Dynamic model

A representative farmer chooses hours of power use, subject to the ration, in order to maximize profits each period. The farmer's problem is

$$\max_{H_t \leq \overline{H}} \Omega(W_t(H_t, D_t))^{\alpha_W} - p_E P H_t. \tag{4}$$

Power use yields water input via the extraction function

$$W_t(H_t, D_t) = \rho \frac{P H_t}{D_t}. \tag{5}$$

The farmer's constrained optimal power and water use are then

$$\begin{aligned} H_t^* &= \min\left\{\left(\frac{\Omega\alpha_W}{p_E}\right)^{\frac{1}{1-\alpha_W}}\left(\frac{\rho}{D_t}\right)^{\frac{\alpha_W}{1-\alpha_W}}\frac{1}{P}, \overline{H}\right\}, \\ W_t^* &= \rho\frac{PH_t^*}{D_t}. \end{aligned} \tag{6}$$

Extracting water today lowers the water level tomorrow. The groundwater law of motion is

$$D_{t+1} = D_t + \gamma\left(W_t - R\right) \tag{7}$$

where W_t is water use and R denotes the recharge rate. Recharge is exogenous and depends on rainfall and geological factors.

Social surplus consists of the present value of farmer profits less the cost the state incurs in supplying power

$$S\left(D_t\right) = \sum_{t=0}^{\infty}\beta^t\left[\Pi(W_t(H_t^*(D_t), D_t)) - (c_E - p_E)PH_t^*(D_t)\right]. \tag{8}$$

Surplus is deterministic given the initial condition D_t, the farmer's constrained input use (6) in each period and the groundwater law of motion (7). The opportunity cost of water is the change in future surplus with respect to a change in water extraction today. Increasing W_t by one unit increases tomorrow's depth by γ and thereby the future path of depth. Hence the opportunity cost of a unit of water extraction is

$$\lambda_W = \frac{dS(D_{t+1})}{dD_{t+1}}\frac{dD_{t+1}}{dW_t} = \frac{dS(D_{t+1})}{dD_{t+1}}\gamma. \tag{9}$$

We calculate this opportunity cost numerically with a finite difference approximation.

b Estimation of dynamic model

There are three sets of parameters to estimate, for the production function, the extraction function and the law of motion. The production function and extraction function parameters have already been estimated in the main text. Table E12 summarizes their values. We consider a representative farmer who has the average productivity from our estimates (inclusive of the effects of non-water inputs, taken as exogenous) and the average well depth and pump capacity.

We estimate the groundwater law of motion (7) by fitting our model to changes in well depth

Table E12
Parameters used in the dynamic model

Parameter	Value	Source
		Primitives
α_W	0.18	Main model
Ω	13.00	Main model
		Exogenous variables
p_E	INR 0.9	Rajasthan policy
c_E	INR 6.2	Rajasthan policy
$\overline{H}$	6 hours	Rajasthan policy

This table reports the inputs to our model that are homogenous across all SDOs.The *primitives* are unobserved structural parameters assumed to be policy invariant.These include α_W, which defines the concavity of the production function,and Ω which is total factor productivity.The *exogenous variables* are unmodeled policy choices which include the nominal price of one kilowatt-hour of electricity,the marginal cost of producing one kilowatt-hour of electricity, and the power ration in hours per day.

for wells drilled at different times. Groundwater extraction in Rajasthan has been lowering the water table year by year, so that farmers who are drilling a fresh well generally go deeper than the average of existing wells. We observe that the depth of new wells has been trending deeper over time at a fairly steady pace for twenty-five years (Figure E10). We take this decline in depth as a proxy for the decline in water levels.

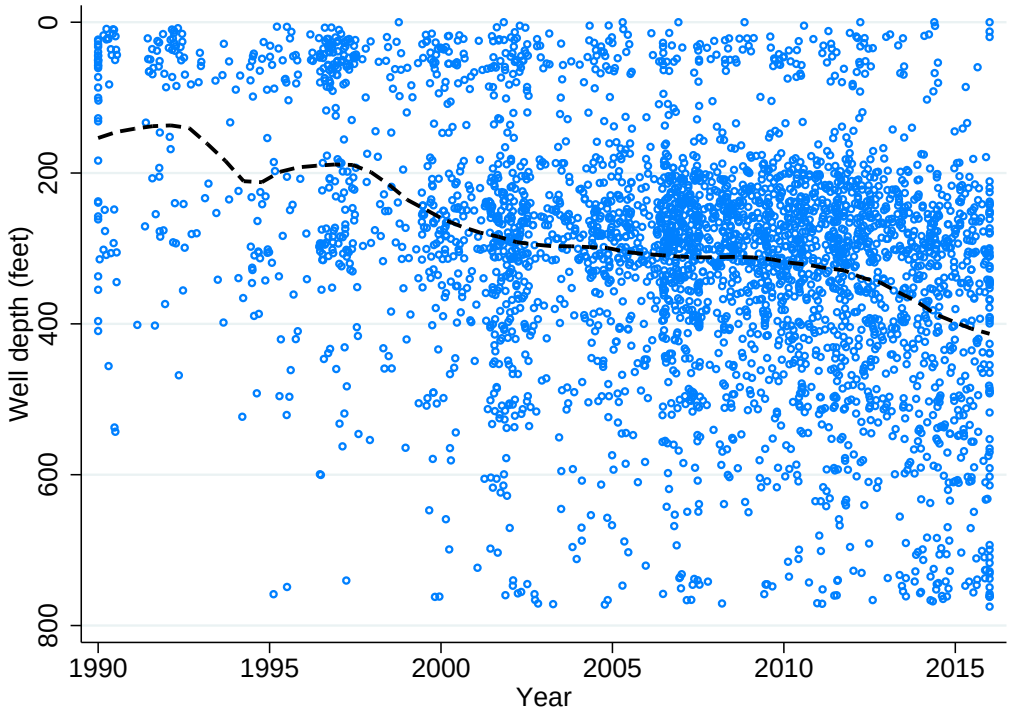
The key parameter in the groundwater law of motion is γ, the effect of water use in a given year on depth in the following year. We estimate γ by finding the value that best matches the observed trend in water depletion in our sample. The procedure is as follows:

1. *Set initial conditions.* We calculate present water use given the terminal depth. We solve the model given the depth of wells drilled in the most recent year to calculate water use. We fix a constant R for recharge. The Government of India estimates the ratio of water extraction to natural recharge, $\delta = \frac{W}{R}$. We use the state-level ratio for Rajasthan of $\delta = 1.4$ to infer the recharge rate from present water input use.

2. *Project depth backwards.* For a given candidate γ and water use, we project the path of well depths backwards using farmer's water input choice at each period and the law of motion.

3. *Optimize over γ.* Our estimate of $\hat{\gamma}$ is then chosen to minimize the sum of squared differences

Figure E10: Depths of wells dug by year



This figure shows the distribution of depths of wells dug by farmers in our sample between the years 1990 and 2016.

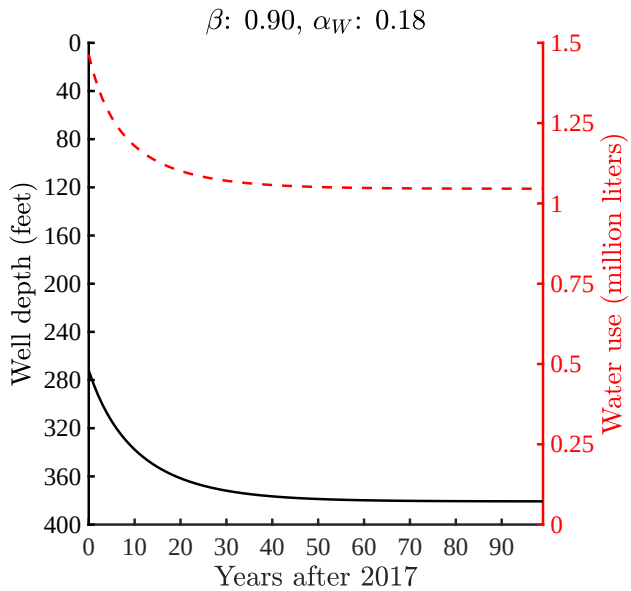
between the model projected well depth in a given year and the actual depth of wells that farmers drilled in that year.

c Results

We estimate the key parameter of the groundwater law of motion to be $\hat{\gamma} = 0.026$ feet per liter (standard error of 0.003 feet per liter). Since our dynamic model has a representative farmer and the decline in depth is estimated based upon that farmer's water use, this parameter represents the change in future water depth if all farmers increased their average water use by a given quantity.

Figure E11 shows projections of well depth and water use from our dynamic model. This projection assumes that the six hour ration will remain in place. The representative farmer will then use either the full ration or a lesser amount as described by (6). We use the path of well depth to represent the path of water depth because there is a tight, linear relationship between these variables, but our data on well depth has broader coverage (Appendix A). Water use affects groundwater depth and therefore well depth via the groundwater law of motion. The solid (black)

Figure E11: Model projections of water use and groundwater depth



This figure shows projected groundwater depth and water use for a representative farmer in the dynamic model. The horizontal axis depicts the number of years which have passed since 2017, the year of our sample data. The left vertical axis (colored black) shows the depth of groundwater over time. The right vertical axis (colored red) shows the corresponding water use. The starting well depth is taken to be the average well depth of all farmers in our sample, which is roughly 272 feet. We use our point estimate of the law of motion $\hat{\gamma} = 0.026$, along with an output elasticity of 0.18 and a discount rate of 0.9 to parameterize the model.

line, measured against the left axis, shows well depth over time. The dashed (red) line, measured against the right axis, shows water use. As these series are for a representative farmer they can be compared to average usage at present.

Figure E11 shows that well depth is projected to increase from about 280 feet today to approximately 380 feet over the next forty years. Many individual farmers today already have wells that are deeper than this predicted representative future depth. By way of comparison, Figure E10 shows that the average depth of new wells in our sample has declined by a greater amount, about 200 feet, over the last thirty years. The pace of the decline in water levels is projected to slow over time, despite a constant policy environment, because farmers are able to extract less water for a given duration of pumping as the water level declines. Eventually, the water table reaches a steady state, where the amount of water extracted is counterbalanced by natural recharge. Our model forecasts that water use will decline from 1.4 million liters per farmer to just over 1 million liters per farmer at that steady state.

Water extraction today, in this way, increases extraction costs tomorrow by lowering the water

Table E13
Estimates of λ_W for alternate parameter values

$\beta \backslash \alpha_W$	0.12	0.15	0.18	0.21	0.24
0.95	1.93 (0.08)	2.98 (0.13)	4.43 (0.19)	6.41 (0.28)	9.08 (0.39)
0.90	1.41 (0.09)	2.18 (0.14)	3.25 (0.21)	4.70 (0.30)	6.66 (0.42)
0.75	0.72 (0.06)	1.11 (0.10)	1.66 (0.15)	2.40 (0.22)	3.41 (0.31)

This table reports the opportunity cost of water for different values of the output elasticity of water α_W and the discount rate β. The units of λ_W are INR per liter. Bootstrapped standard errors in parentheses account for estimation error in the groundwater law of motion.

table. Table E13 reports our estimates of the opportunity cost of water. We calculate the opportunity cost of water for a range of values of the output elasticity of water α_W (across columns of the table) and the discount rate β (across rows). We estimate the water elasticity α_W as part of the production function in Table 3. For the discount factor, we consider several values meant to capture borrowing costs for the state or for farmers themselves. Our main estimates use a discount factor of $\beta = 0.90$, which is close to one less the nominal interest rate on Rajasthan's state government bonds. We also consider a higher discount factor of $\beta = 0.95$, which is closer to the real rate of interest on state bonds, and a lower discount factor of $\beta = 0.75$. We expect that the interest rates faced by farmers in their own borrowing will generally exceed $0.25 = 1 - 0.75$.

Our focal estimate of the value of λ_w is INR 3.25 per thousand liters, which we use in the main text and counterfactual results. As expected, higher discount factors, or higher elasticities of output with respect to water, both increase the estimated value of water. With our estimated value of $\alpha_w = 0.18$ (Table 3, column 4) and the higher discount factor $\beta = 0.95$, the opportunity cost of water increases to INR 4.57 per thousand liters (36% higher); at the lower discount factor of $\beta = 0.75$ the opportunity cost of water is INR 1.71 per thousand liters (49% lower). Since the social cost of power use is about evenly split between the private cost of power supply and the opportunity cost of the water extracted, the same changes in the discount factor have proportional effects on the social cost of water extraction that are only about half as large as their effects on the opportunity cost component (λ_w) alone.

References

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen. 2012. "Sparse models and methods for optimal instruments with an application to eminent domain." *Econometrica*, 80(6): 2369–2429.

Blakeslee, David, Ram Fishman, and Veena Srinivasan. 2020. "Way down in the hole: adaptation to long-term water loss in rural India." *American Economic Review*, 110(1): 200–224.

Gollin, Douglas, and Christopher Udry. 2020. "Heterogeneity, Measurement, and Misallocation: Evidence from African Agriculture." *forthcoming, Journal of Political Economy.*

Jasmin, Ismail, and P Mallikarjuna. 2011. "Satellite-based remote sensing and geographic information systems and their application in the assessment of groundwater potential, with particular reference to India." *Hydrogeology Journal*, 19(4): 729–740.

Lee, S, I Park, and JK Choi. 2012. "Spatial Prediction of Ground Subsidence Susceptibility Using An Artificial Neural Network." *Environmental Management*, 347–358.

Lockhart, Richard, Jonathan Taylor, Ryan J Tibshirani, and Robert Tibshirani. 2014. "A significance test for the lasso." *Annals of statistics*, 42(2): 413.

Mallick, Javed, Chander Kumar Singh, Hussein Al-Wadi, Mohd Ahmed, Atiqur Rahman, Satyanarayan Shashtri, and Saumitra Mukherjee. 2015. "Geospatial and geostatistical approach for groundwater potential zone delineation." *Hydrological Processes*, 29(3): 395–418.

Manring, Noah D. 2013. *Fluid Power Pumps and Motors: Analysis, Design, and Control.* McGraw-Hill Education.

Oxford University Press. 2011. "India Infrastructure Report." Oxford University Press, Oxford University Press.

Pakhale, Gaurav, Gupta Prasun, and Jyoti Nale. 2010. "Crop And Irrigation Water Requirement Estimation By Remote Sensing And GIS: A Case Study Of Karnal District, Haryana, India." *International Journal of Engineering and Technology*, 2: 207–211.

Sander, Per. 2007. "Lineaments in groundwater exploration: a review of applications and limitations." *Hydrogeology Journal*, 15(1): 71–74.

Shakti Foundation. 2016. "Case Studies of DSM Projects Implemented in India." Shakti Foundation, Shakti Foundation.

Timmins, Christopher. 2002. "Measuring the dynamic efficiency costs of regulators' preferences: Municipal water utilities in the arid west." *Econometrica*, 70(2): 603–629.