Markets, Contracts, and Uncertainty: A Structural Model of a Groundwater Economy

Xavier Giné and Hanan Jacoby*

June 11, 2014

preliminary draft; please do not cite

Abstract

We develop a contract-theoretical model of groundwater transactions in rural India, emphasizing the role of payoff uncertainty, which arises in our context due to unpredictable fluctuations in groundwater availability during the dry season. Our focus is on the tradeoff between the ex-post inefficiency of seasonal contracts and the ex-ante inefficiency of more flexible water-selling arrangements. We structurally estimate the model using micro-data on area irrigated under each transaction type combined with subjective probability distributions of end-of-season well output collected from over 2000 well-owners across six districts of Andhra Pradesh. The estimated parameters are used to back out the implicit cost of excessive land fragmentation, the absence of which would render groundwater markets superfluous.

1 Introduction

Uncertainty is often invoked to rationalize market institutions in developing countries. Phenomena like share-tenancy (Stiglitz, 1974) and intra-village risk-sharing (Townsend, 1994), to cite but two prominent strands of the literature, have been seen as (informal) contractual adaptations by risk-averse agents to risky environments. In this paper, we consider a novel channel through which uncertainty shapes the character and reach of market institutions, one that does not rely on risk aversion. We examine theoretically and evaluate quantitatively the role of payoff uncertainty – the variability of quasi-rents in a trading relationship with respect to the state of the world (Hart, 2009).

The context for our investigation is the groundwater economy of rural Andhra Pradesh. As in much of South Asia, dry-season agricultural production in this large south Indian state relies almost exclusively on irrigation water extracted from private borewells. Yet, the supply of groundwater throughout the season is uncertain. Key features of this uncertainty are, firstly, that it resolves only after planting investments have been made and, secondly, that it affects the (marginal) returns to these investments. Given the irreversibility of cultivation decisions, well-owners may evince a precautionary motive analogous to that in the savings literature (e.g., Kimball, 1990), in this case limiting their exposure to supply uncertainty by committing less area to irrigate in the dry-season.

Against this backdrop, we study water-selling institutions. The two principal arrangements are seasonal contracts, which specify quantity and price of irrigation ex-ante, and per-irrigation sales, which are negotiated throughout the course of the season. Drawing on the recent insights of Hart and Moore (2008) and, especially, Hart (2009), we view these institutions as trading off ex-post against ex-ante inefficiency. In particular, the seasonal contract, by establishing an entitlement, makes renegotiation or hold-up costly, prohibitively so in our set-up, but limits flexibility once payoff uncertainty is resolved. By contrast, per-irrigation sales, being the outcome of bargaining, distort ex-ante planting investments, but achieve an efficient groundwater allocation ex-post. A key theoretical prediction is, therefore, that as payoff uncertainty increases, water-sellers will replace seasonal contracts with per-irrigation arrangements. Moreover, since either type of transaction entails a distortion, groundwater markets overall will be less developed where uncertainty is greater.¹

¹Jacoby et al. (2004) find another type of groundwater market distortion in Pakistan: Monopolistic well-owners price water above its marginal extraction cost. They speculate that fear of hold-up precludes the type of seasonal contracts (i.e., with lump-sum transfers) that would achieve an efficient allocation. Banerji et al. (2011) also consider groundwater markets in South Asia within a bargaining framework.
Prior to Hart and Moore’s (2008) idea that contracts act as reference-points, payoff uncertainty was not thought to have any bearing on contractual form. In the standard theory of hold-up (e.g., Grossman and Hart, 1986; Hart and Moore, 1990), renegotiation is efficient, which is to say that no surplus is destroyed in the process of ex-post haggling; under these conditions, hold-up is virtually inevitable (Hart, 1995). While lab experiments have established support for the reference-point theory (Fehr et al., 2011; Hoppe and Schmitz, 2011), to our knowledge the present paper is the first application to a real-world contracting environment.

We formalize the abovementioned precautionary planting motive and the tradeoff between rigid and flexible water-selling arrangements within a unified model of agricultural production under stochastic groundwater supply. Thus, our structural model accounts for alternative modes of water-selling, as well as for water transfers through leasing, and, given specific parameterizations of the crop production and cost functions, delivers expressions for the optimal amount of land irrigated under each such arrangement. We estimate the model by maximum likelihood on a large sample of borewells covered in a specially-designed groundwater markets survey undertaken across six districts of Andhra Pradesh. The survey takes particular care to elicit from each well-owner their subjective probability distribution of groundwater supply near the end of the season. The structural parameters are identified principally off of variation across borewells in this probability distribution.2

Within the empirical contracts literature, Gagnepain et al. (2013) is perhaps most relevant inasmuch as they also consider the tradeoff between ex-post renegotiation and ex-ante incentives, in the context of French public-sector contracts, and devise a structural econometric model to quantify the welfare costs of limited commitment. In their case, however, contractual choice (between fixed-price and cost-plus) is driven by asymmetric information rather than payoff uncertainty. Moreover, in our setting, agents have the option not to contract or trade at all (and many do not), which allows us to investigate the extent to which payoff uncertainty limits the development of markets.

Using the structural parameter estimates, we perform two types of exercises. First, we simulate the response of irrigated area, overall and under each specific transfer arrangement, to changes in uncertainty. We then go on to ask whether, or rather how well, variation in uncertainty can account for the stark differences in groundwater markets across our six sam-

---

2While several recent papers incorporate subjective probabilities into structural econometric models (see, e.g., Delvande, 2008; Mahajan et al., 2011; Mahajan and Tarozzi, 2012), ours is the first such application in the contracts literature. Delavande et al. (2010) discuss issues in collecting subjective expectations data in developing countries.
ple districts. Second, we evaluate the economic gains from a hypothetical land consolidation program by estimating how much more area would be irrigated if all plots surrounding a given well were owned by the same individual.

The road-map for the remainder of this paper is as follows: The next section describes the survey data and the groundwater economy of Andhra Pradesh. Section 3 lays out the theoretical model and derives comparative statics results for changes in uncertainty. Section 4 discusses the empirical specification including a derivation of the likelihood function. Estimation results and counterfactual simulations are reported in Section 5. Section 6 concludes.

2 Data and Context

Our data come from a randomly selected survey of about 2300 borewell owners undertaken in six districts of Andhra Pradesh (AP) in 2012-13. The districts were selected to cover a broad range of groundwater availability, conditions for which generally improve as one moves from the relatively arid interior of the state toward the lusher coast. Drought-prone Anantapur and Mahbubnagar districts were originally selected as part of a weather-index insurance experiment (Cole et al., 2013); all 750 or so borewell owners were followed up from that study’s 2010 household survey. Kadapa and Guntur districts, which fall in the intermediate range of rainfall scarcity, and the water-abundant coastal districts of East and West Godavari, each contribute around 400 borewell owners to our sample.³

2.1 Recharge and uncertainty

As in much of India, farmers in AP rely almost exclusively on groundwater during the *rabi* (winter) season, when rainfall is minimal and surface irrigation typically unavailable. Indeed, the last two decades have seen an explosion of borewell investment as the cost of drilling and of submersible electric pumpsets have fallen.⁴ Despite alarm about groundwater overexploitation in India more broadly (e.g., *New York Times*, 2006; *Economist*, 2009), water-tables across AP do not exhibit much downward trend; rather, the time-series is dominated by inter-annual variability (Appendix figure A.2). This has to do with the limited storage capacity of the shallow hard rock aquifers that characterize the region. Most

³A total of 144 villages were covered (21-25 per district) in the survey. Our sample is broadly representative of areas where sufficient groundwater can be had for *rabi* cultivation and where groundwater is the sole source of irrigation during that season; i.e., we explicitly avoided villages in canal command areas.

⁴Appendix figure A.1 documents the rising importance of borewell irrigation in all of India, in AP as a whole, and in the six districts of AP selected for our survey. See also World Bank (2005).
of the recharge from monsoon rains occurring over the summer months is depleted through groundwater extraction in the ensuing rabi season. In contrast to the hydrogeology of much of North India, there are no deep groundwater reserves to mine (see Fishman et al., 2011).

This annual cycle of groundwater replenishment and draw-down throughout AP is one of two crucial ingredients to our story. Even though farmers can observe monsoon rainfall along with their own borewell’s discharge prior to rabi planting, they cannot perfectly forecast groundwater availability over the entire season. To assess the degree of uncertainty, we fielded a well-flow expectations module as part of the borewell owner’s survey, which was structured as follows: First, we asked owners to assess the probability distribution of flow on a typical day at the start of (any) rabi season, the metric for discharge being fullness of the outlet pipe (i.e., full, \( \frac{3}{4} \) full, \( \frac{1}{2} \) full, \( \frac{1}{4} \) full, empty). Next, using the same format, we asked about the probability distribution of end-of-season flow conditional on the most probable start-of-season flow. Thus, the question was designed to elicit residual uncertainty about groundwater availability. Figure 1 shows the distributions of groundwater uncertainty (well-specific coefficients of variation of end-of-season flow) across the six districts of our study. The first thing to notice is that virtually no one reports having a perfectly certain supply of groundwater. Moreover, the pattern of variation across districts is suggestive: higher uncertainty where aquifer recharge is meager (Anantapur and Mahbubnagar); lower uncertainty where recharge is ample (East and West Godavari).

2.2 Land fragmentation, fixed costs, and groundwater markets

The second crucial ingredient to our story is land fragmentation coupled with the high fixed cost of borewell installation, on the order of US$1000 (excluding the pump-set). Fragmentation is driven by the pervasive inheritance norm dictating equal division of land among sons and the prohibitive transaction costs entailed in consolidating spatially dispersed plots through the market.\(^5\)

None of this would matter, of course, if groundwater markets were frictionless. In this case, borewells would be just as likely on small plots as on large plots; the owner of a small plot could simply sell any excess groundwater to a neighbor. Obversely, small plots would be just as likely cultivated in the dry season as large plots; any plot owner without a borewell of his own could purchase groundwater from that of a neighbor. Neither implication of frictionless groundwater markets, however, is consistent with our data.

\(^5\)Appendix figure A.3 illustrates the increasing fragmentation in India, and in AP, as seen through the rising proportion of marginal farms (those with operational landholdings of less than 1 hectare).
Our survey covers around 9600 plots, each of which either has a borewell itself or is adjacent to a plot that does and, thus, can receive a transfer of groundwater in principle. Figure 2 shows that borewells are actually much less likely on small plots than on large ones. To be sure, owners of small plots may also tend to have small landholdings overall and, hence, low wealth. Such farmers may, therefore, be less able to mobilize the resources for borewell installation (see, e.g., Fafchamps and Pender, 1996). To control for this, we focus only on the subset of plots whose owner has at least one other plot (otherwise, plot area and total owned area are perfectly correlated). After partialling out the effect of wealth (as proxied by landholdings) using dummies for each of the deciles of total landownership in figure 3, we continue to find borewell installation increasingly probable as plot size increases. Finally, figure 4 shows that small plots are much more likely to be left fallow in the dry season than large plots. Taken together, this evidence emphatically points to frictions in groundwater markets, the precise source of which are the topic of the remainder of this paper.

2.3 Adjacency format

To capture transfers of groundwater, which typically occur between adjacent plots so as to minimize conveyance losses, we designed a novel questionnaire. Each respondent (borewell owner) was asked to provide information about all the plots adjacent to the one containing the reference borewell, including characteristics of the landowner, details on how the plot was irrigated during the rabi, if not left fallow, and on the transfer arrangement if one occurred. This ‘adjacency’ format has the advantage of telling us not only about the transfers that did happen but also about those that potentially could have happened but did not.

In figure 5, we aggregate plot-specific data up to the level of the roughly 2300 adjacencies. The top curve, which shows how borewell density in the adjacency (the average of the plot-level borewell indicator used in figure 2 weighted by the plot’s area share in the adjacency) varies with average plot size, virtually replicates figure 2 at the adjacency level; borewell development is less intensive on highly fragmented land. The bottom curve refers to the proportion of plots in the adjacency receiving any groundwater transfer from the reference well (aside from transfers between well co-owners). The relationship between this groundwater transactions frequency and average plot size in the adjacency is essentially a

---

6 Most irrigation water is transferred through unlined field channels with high seepage rates. However, our survey also picked up a number of transfers to non-adjacent plots using PVC pipe. Usually, these cases involved sharing of groundwater between well co-owners or between multiple plots of the same owner.

7 To be precise, 2307 borewell owners were interviewed, but there were 114 cases where the reference plot had a second borewell and 2 cases where it had a third borewell, giving a total of 2423 reference wells.
mirror-image of that between borewell density and average plot size.\(^8\) Evidently, borewell density and groundwater market activity are substitutes, both driven by the degree of land fragmentation.

### 2.4 Precautionary planting and risk aversion

As we show in the next section, uncertainty in groundwater availability interacts with the irreversibility of the planting investment in an important way. However, the empirical implications of this interaction turn out to depend critically on the form of the production function, which is not directly observable. To motivate our modelling choices, therefore, we now present a reduced-form analysis of planting decisions.

* \(Rabi\) season cultivation in AP falls into two broad categories, wet crops (in our six districts, principally paddy, banana, sugarcane, and mulberry) and irrigated-dry or ID crops (e.g., groundnut, maize, cotton, chillies), distinguished by the much greater water requirements of the former. In particular, a field that if planted to ID crops would take 3 days to irrigate, would, if under wet crops, take a week to irrigate. Since treating these crop-types separately in either the theory or the empirics would create substantial complications (see below) without providing much in the way of additional insight, we henceforth club them together using the equivalence 1 acre wet = \(\frac{7}{3}\) acre ID.

We investigate the effect of uncertainty on planting decisions using both subjective and objective outcome measures. The subjective outcome \(Y_{is}^s\) for borewell \(i\) is derived from a question in our survey asking each owner how many acres (wet and/or ID) would have been irrigated with his borewell “if at the start of the \(rabi\) you knew for certain what the flow during the rest of the season was going to be.”\(^9\) Thus,

\[
Y_{is}^s = \log(\text{actual irrigated area}) - \log(\text{irrigated area under certainty}).
\]

Based on the sample of about 2400 borewells, \(Y_{is}^s = -0.280\) (median = \(-0.231\)), indicating a (perceived) reduction in irrigated area due to uncertainty on the order of 25%. To see how such precautionary planting behavior varies with groundwater uncertainty, we report in panel (a) of Table 1 a regression of \(Y_{is}^s\) on the \(\log(CV_i)\) of end-of-season flow (see figure 1), as well

---

\(^8\)Each of the local polynomial fits in the figure retain practically identical shapes after adjusting out reference plot/well characteristics (plot area, area-squared, and outlet pipe diameter).

\(^9\)The question refers specifically to the most recently completed \(rabi\) (2011-12). For proper framing, the immediately prior question reestablishes the total area actually irrigated by the same well during the last \(rabi\).
as the log of outlet pipe-width (to control for borewell capacity). The negative coefficient on $\log(CV_i)$ in column 1 indicates that, consistent with a precautionary motive, borewell owners reduce their irrigation commitments in response to higher supply uncertainty. Even after augmenting the baseline specification, in columns 2 and 3, with, respectively, additional well characteristics (pump horse-power, log well depth, number of nearby wells, presence of groundwater recharge) and district fixed effects, our main finding persists.

Next, we consider,

$$Y^o_i = \log(\text{actual irrigated area}) - \log(\text{reference plot area}),$$

replacing the borewell owner’s subjective perfect-certainty counterfactual with the autarky benchmark (i.e., irrigating exactly the borewell-plot area). Thus, a value of $Y^o_i$ less than zero (27% of cases) indicates that part of the borewell’s plot was left fallow in the past *rabi* season, whereas a value greater than zero (46% of cases) indicates that groundwater was either sold (irrigating the land of another farmer in the adjacency) or was transferred to a leased plot. Columns 4-6 in panel (a) of Table 1 report for $Y^o_i$ the analogous set of regressions we ran for $Y^*_i$, with qualitatively similar results.\(^{10}\)

While the evidence, thus, strongly supports a precautionary planting motive, we have not yet established the theoretical mechanism. To the extent that variability in irrigation supply induces fluctuations in household income, simple risk aversion may explain why farmers limit their *rabi* planting in the face of uncertainty. To assess this, we use two measures of preferences towards risk collected from well owners in the survey. The first measure, $RISK^1_i$, is a self-assessed ranking of risk tolerance, with 1 indicating “I am fully prepared to take risks” and 10 indicating “I always try to avoid taking risk.” The second measure is based on a Binswanger (1980) lottery played by each respondent for real money. Following Cole et al. (2013), $RISK^2_i$ is an index of marginal willingness-to-pay for risk constructed from the characteristics of the preferred lottery. Ranging from 0 to 1, higher values of $RISK^2_i$ indicate greater risk aversion. Panels (b) and (c) of Table 1 report results of adding, respectively, $RISK^1_i$ and $RISK^2_i$, and, most importantly, their interactions with $\log(CV_i)$, in the corresponding baseline regression of panel (a). The estimated coefficients on these interactions, and particularly their lack of significance, betrays no indication that highly risk

\(^{10}\)One concern is the following: Suppose that actual irrigated area is, in fact, unresponsive to groundwater uncertainty but $\log(\text{reference plot area})$ and $\log(CV_i)$ happen to be positively correlated. In this scenario, we would find a negative association between $Y^o_i$ and $\log(CV_i)$, which is purely mechanical. It turns out, however, that $\log(\text{reference plot area})$ is *negatively* correlated with $\log(CV_i)$, whether or not we condition on the other regressors.
averse borewell owners are especially responsive to groundwater uncertainty. Precautionary planting, therefore, does not appear driven by risk preference.

3 Theoretical Framework

3.1 Model assumptions

Our basic theoretical assumptions are as follows:

A.1) **Fragmentation:** Agricultural production occurs on discrete plots of land of area $a_i$, each owned by a distinct individual.

As mentioned, fragmentation presupposes that the outright purchase of neighboring plots is generally infeasible, perhaps due to a combination of mortgage finance constraints and thinness of the land market.

A.2) **Borewells and groundwater:** A reference plot has a borewell drawing a quantity of groundwater $w_i$ over the growing season, where $w_i$ has a discrete distribution $H(w_i)$ with $k = 1, 2, ..., K$ points of support $w_k$ and corresponding probabilities $\pi_k$. Let $\bar{w} \equiv E[w] = \sum_k \pi_k w_k$ and $\sigma_w^2 \equiv Var[w]$.

Our focus is on the allocation of water across plots from existing wells. Thus, we abstract from the decisions to explore for groundwater and to invest in borewells. Because tariffs for agricultural power are zero in AP, farmers pump the maximum they can from their borewells given the fixed daily electricity ration. Aside from this power constraint, $w_i$ depends on the capacity of the well (pipe-width), the availability of groundwater in the aquifer (i.e., recharge), and the local hydrogeology.

A.3) **Crop production:** $y_i/l = f(w_i/l) \equiv f(\omega)$, where $y_i$ is crop output, $w_i$ is assumed to be the only water source, $l_i$ is the area irrigated, $\omega$ is irrigation intensity, and $f$ is an increasing, concave, function with $f(0) = 0$.

Constant returns to scale is a sensible assumption for reasons both technical and empirical. As to the latter, diminishing returns is unlikely to set in over the range of cultivated areas that we are considering. Moreover, under diminishing returns, well-owners might simultaneously
leave their own plot partially fallow while selling water to a neighboring plot, a scenario which is virtually never observed in practice.\footnote{Constant returns also imposes an identifying restriction, leaving one less parameter (the output elasticity with respect to land area) to be estimated.}

Also embedded in (A.3), is our convention (noted earlier) that \( l \) is expressed in \textit{efficiency units} of ID crop area; i.e., area under wet crops is converted into area under ID at the fixed rate of 7:3. In this sense, we do not take a stand on the choice between these two crop-types; farmers are assumed to be indifferent. There are three practical obstacles to explicitly modelling the cultivation choice between separate wet and ID crop technologies: (1) Aside from the extra margin of ex-ante area choice, farmers would also allocate groundwater ex-post across crops in response to the realized \( w \), leading to one additional optimality condition for each state of nature; (2) To rationalize ID or wet mono-culture, the structural model would have to account for corner solutions in cropped area; (3) For any form of groundwater transfer, each cell of the \( 3 \times 3 \) matrix of wet-ID-mixed cropping decisions of borewell owner and groundwater recipient would have to be compared to determine the optimal arrangement. As already remarked, these technical complexities and the fact that crop choice is tangential to our primary focus militate in favor of the efficiency units assumption.

A.4) \textit{Cultivation costs:} There is a constant marginal cost \( c \), in terms of output, per unit of land cultivated so that net revenue is \( y - cl \).

For theoretical consistency, \( c \) incorporates the opportunity cost of land rental, even for farmers who cultivate their own land.

And, finally, as alluded to at the end of the last section

A.5) \textit{Risk preferences:} Farmers/well-owners are risk neutral.

### 3.2 Unconstrained cultivation choice

Consider a well-owner's choice of area cultivated (irrigated) when his own plot size is not a constraint. Let \( l_\sigma = \arg \max l \{Ef(w/l) - c\} \) and \( l_0 \) be the optimal area cultivated under certainty; i.e., \( l_0 = l_\sigma \) when \( \sigma_w^2 = 0 \). The following function will appear repeatedly in the sequel

\textbf{Definition 1} \( g(\omega) = f(\omega) - \omega f'(\omega) \).

We can now prove
Proposition 1 (Precautionary planting) If \( g \) is increasing and strictly concave, then \( l_\sigma < l_0 \).

Proof. The necessary condition for an optimum is \( E g(w/l_\sigma) = c \), whereas under certainty it is \( g(\bar{w}/l_0) = c \). Thus, \( g(\bar{w}/l_0) = E g(w/l_\sigma) < g(\bar{w}/l_\sigma) \), where the latter inequality follows from the concavity of \( g \).\(^{12}\) The result follows from the assumption that \( g \) is increasing.

For consistency with subsequent notation, let \( \ell_S = a - l_\sigma \) be the area of the plot left fallow if \( l_\sigma < a \). The surplus from unconstrained self-cultivation is as follows:

Definition 2 \( V_S = l_\sigma E[f(w/l_\sigma) - c] \)

In case \( l_\sigma > a \), we may think of \( V_S \) as the surplus derived by the well-owner if he could sell groundwater in a competitive spot market.

3.3 Leasing

If a well-owner’s plot size does constrain his cultivation, he could lease in adjacent land. We assume, however, that leasing entails an efficiency cost that makes it less attractive than own-cultivation. One rationalization for such costs, corroborated by Jacoby and Mansuri (2009), is that underprovision of non-contractible investment (e.g., soil improvement) lowers the productivity of leased land. At any rate, without invoking some sort of leasing cost, the existence of a market for groundwater and, indeed, the predominance of groundwater sales over leasing would be problematic.

Thus, letting \( \gamma \) be a percentage increment to cultivation costs, define

\[
\ell_L = \arg \max (a + \ell) \{ Ef(w/(a + \ell)) - c \} - \gamma c \ell.
\]

As should be clear from this expression, the efficiency cost only applies to leased land.

One asymmetry in our approach worth noting at the outset, because it will be common across all water transfer arrangements, is that, even though the well-owner is constrained by his own plot size, there is no such constraint on the amount of adjacent land available for irrigation, in this case through leasing. While this assumption vastly simplifies the model, it is clearly not realistic and, for this reason, we will allow for supply-side heterogeneity in the empirical work.

\(^{12}\)If the coefficient of relative prudence (Kimball, 1990), \(-\omega f'''/f''\), exceeds unity, then \( g \) is concave. Thus, the precautionary planting motive hinges on the convexity of the marginal product of water, \( f' \), just as the precautionary savings motive hinges on the convexity of the marginal utility of consumption. If, for example, \( f \) is quadratic (linear marginal product), then \( g'' < 0 \). In this case, more land is cultivated under uncertainty than under certainty.
3.4 Seasonal water contracts

The canonical water contract in our setting commits the well-owner to irrigate a buyer’s field, or some portion thereof, for the whole season at a pre-determined price. Following Hart and Moore (2008) and Hart (2009), we think of such contracts as establishing an entitlement. Ex-post renegotiation of the terms, or hold-up, would therefore lead to deadweight losses due to aggrievement by one or both parties. To bring the tradeoff between contractual rigidity and flexibility into stark relief, we assume that these deadweight losses make hold-up prohibitively costly and we allow ex-post water trade as a non-contractual alternative. Thus, the seasonal contract has two salient features: (1) by serving as reference point in, and hence as a deterrent to, renegotiation, it protects relationship-specific investment (in our context, planting inputs); and (2) water allocations under the contract are unresponsive to the state of the world, a distortion which necessarily increases with uncertainty.

Let $\tau$ denote the total transfer of irrigation water at a per unit price of $p$ applied to a field of size $\ell$. Assuming that all farmers share the same technology, the optimal simple (i.e. single price) contract solves

$$\max_{p, \ell} a \left\{ Ef\left(\frac{w - \tau}{a}\right) - c \right\} + p\tau \quad \text{s.t.}$$

$$PC : \ell \left\{ f\left(\frac{\tau}{\ell}\right) - c \right\} - p\tau \geq 0$$

$$IC : \tau = \arg\max_{\tau \in (0, w_1]} \ell \left\{ f\left(\frac{\tau}{\ell}\right) - c \right\} - p\tau$$

where the expectations in the first line are taken over $H$. Note that expectations are dropped in the participation constraint ($PC$) and incentive constraint ($IC$) because both $\ell$ and $\tau$ are fixed ex-ante, the latter subject to the constraint that the promised transfer cannot exceed

---

13Hart (2009) argues that a contract indexed to a signal perfectly correlated with the state of the world would achieve the first-best. In our setting, an obvious form of indexing would be a share-contract whereby the buyer delivers some fraction of his harvest in exchange for water. Yet, we do not observe share-contracts for groundwater in our data (only three odd cases). Perhaps, the well-established disincentive effects of share-contracts (e.g., Shaban, 1987) render such arrangements unattractive.

14More precisely, these papers assume that there are noncontractible actions that either party can take ex-post to add value to the transaction. As long as a party feels he is getting what he is entitled to in the contract, he will undertake such helpful actions, but if he feels shortchanged he will withhold them, generating a loss in surplus. In the words of Hart (2009): “Although our theory is static, it incorporates something akin to the notion of trust or good will; this is what is destroyed if hold-up occurs.” (p. 270).

15Since Hart (2009) and Hart and Moore (2008) only consider transactions involving a single ‘widget’, ex-post distortions due to quantity misallocation do not arise in their set-up. Instead, the cost of a rigid contract is the reduced likelihood of trade occurring.
the available supply of water in the worst state of the world, \( w_1 \). Thus, the seasonal contract offers an assured supply of irrigation; the direct cost of production variability is borne fully by the water-seller on his plot in exchange for a monetary transfer from the buyer.

Given a binding \( PC \), the necessary conditions for the optimal contract are as follows:

\[
E f'(\frac{w - \tau}{\alpha}) = p
\]
\[
g'(\frac{\tau}{\ell}) = c
\]
\[
f'(\frac{\tau}{\ell}) = p
\]

Divergence of supply and demand for water ex-post creates the distortion; i.e., it is not the case that \( f'(\frac{w-\tau}{a}) = f'(\frac{\tau}{L}) \ \forall \ w \), which would obtain if \( p \) and \( \tau \) were state-contingent, as in a competitive spot market. It follows as a corollary that the distortion vanishes when \( \sigma_w^2 = 0 \), in which case \( g(\frac{\tau}{\ell_C}) = g(\frac{w}{\ell_C+a}) = c = g(\frac{w}{l_a}) \Rightarrow \ell_C = l_a - a = l_0 - a \).

### 3.5 Water trade without contracts

Outside of a seasonal contract, parties may transact water on a per-irrigation basis. Once the season is underway, however, commitments have been made: The potential water-seller has retained (i.e., refrained from contracting out) the rights to some excess water from his well during the season whereas the potential water-buyer has planted a crop in an adjacent plot. Since each party, therefore, has some degree of bargaining power, we use a Nash bargaining framework.

Before proceeding, we should elaborate on what we mean here by absence of contracts. There is, after all, a self-enforcing agreement to trade during the course of the season, even though the terms of these trades are not necessarily delineated ex-ante. Indeed, side-payments may be made (or favors rendered) to secure an exclusive trading relationship. We assume that any negotiations at this stage are efficient; in other words, the parties will leave no money on the table.

Returning to the ex-post stage, let \( \hat{\tau} \) be the amount of water already transferred to the buyer and suppose that buyer and seller negotiate the price \( p \) of incremental transfer \( \Delta \). The buyer’s net payoff from consummating the trade is given by \( u = \ell f((\hat{\tau} + \Delta)/\ell) - p\Delta \), whereas that of the seller is \( v = af((w - \hat{\tau} - \Delta)/a) + p\Delta \). The no-trade payoffs are given by \( u = \ell f(\hat{\tau}/\ell) \) and \( v = af((w - \hat{\tau})/a) \), respectively. The absence of \( c \) in these payoff functions
reflects the fact that all cultivation costs have already been sunk.

Given Nash bargaining, \( p^* = \arg \max (u - v)^\eta (v - u)^{1-\eta} \), where \( \eta \) is the buyer’s bargaining weight. Therefore, \( p^* \) solves

\[
\eta a \left[ \frac{f\left(x^\ell_a - \Delta - x^a\right)}{\Delta} - \frac{f\left(x^\ell_a - x^a\right)}{\Delta} \right] - (1 - \eta)\ell f'\left(\frac{w - \tau - \Delta}{\ell_a}\right) + p = 0
\]

\[
-\eta f'\left(\frac{w - \tau}{a}\right) - (1 - \eta) f'\left(\frac{\tau}{\ell_a}\right) + p = 0
\]

where the last line takes the limit of the second line as \( \Delta \to 0 \). Thus, we obtain the standard surplus-splitting rule

\[
p^*\left(\frac{\tau}{\ell_a}\right) = (1 - \eta) f'\left(\frac{\tau}{\ell_a}\right) + \eta f'\left(\frac{w - \tau}{a}\right).
\]

Furthermore, demand for irrigation dries up once buyer surplus \( f'\left(\frac{\tau}{\ell_a}\right) - p^*\left(\tau\right) = \eta \left[ f'\left(\frac{\tau}{\ell_a}\right) - f'\left(\frac{w - \tau}{a}\right) \right] = 0 \); i.e., when \( \tau \) satisfies \( f'\left(\frac{\tau}{\ell_a}\right) = f'\left(\frac{w - \tau}{a}\right) \).

Now consider the buyer’s ex-ante problem of choosing area cultivated to maximize expected returns given price function \( p^*(\tau) \). In particular,

\[
\ell_P = \arg \max E \left\{ \ell f\left(\frac{\tau}{\ell_a}\right) - \int_0^\tau p^*(t)dt \right\} - c\ell.
\]

Observe that the per unit price of water is no longer constrained to be constant as in the seasonal contract; each small increment of irrigation now has a different cost. Since, from (1), \( \int_0^\tau p^*(t)dt = (1 - \eta)\ell f\left(\frac{\tau}{\ell_a}\right) + \eta a \left[ f\left(\frac{w - \tau}{a}\right) - f\left(\frac{w}{a}\right) \right] \), and only the first term on the right-hand side depends on \( \ell \), the necessary condition for the buyer’s cultivation choice is simply \( \eta Eg(\tau/\ell_P) = c \). Comparing this expression to the case of unconstrained cultivation choice described in Proposition 1, we see that they differ by the factor \( \eta \). Surplus extraction on the part of the water seller effectively taxes the marginal benefits of cultivation, with the tax rate decreasing in the buyer’s bargaining power. Due to this tax-wedge, \( \ell_P < l_a - a \) at \( \sigma_w^2 = 0 \).

While a fuller comparison of the two modes of water selling must await the imposition of additional structure, we can already draw an important conclusion: Whereas the distortion of the seasonal contract disappears when irrigation supply becomes perfectly certain (as we have just seen), the ex-post bargaining distortion does not disappear. The latter inefficiency essentially arises from the hold-up problem first formalized by Grout (1984); the buyer under-
invests in anticipation of ex-post rent appropriation. So, removing uncertainty in water supply does not eliminate this distortion. By contrast, the rigidity of the seasonal contract leads to an inefficient ex-post water allocation only insofar as water supplies are uncertain.

3.6 The role of uncertainty

To make further progress, we add assumption

A.6) Functional form: \( f(\omega) = \zeta \omega^\alpha, \ 0 < \alpha < 1. \)

Henceforth, we normalize \( \zeta = 1 \), which is tantamount (given the homogeneity of \( f \)) to fixing the units of output and, thus, of cultivation costs \( c \). The power function has two desirable properties for our purposes. First, it implies that \( g \) is increasing and globally concave, so that there is always a precautionary planting motive.\(^{16}\) This is at least consistent with the empirical evidence presented in Section 2. Second, the function delivers a closed form solution for area irrigated in most cases.

3.6.1 Unconstrained self-cultivation

As just mentioned, under (A.6), \( l_\sigma < l_0 \) from Proposition 1. In particular, \( Eg(w/l) = (1 - \alpha)E(w/l)^\alpha = c \) implies

\[
\begin{align*}
l_\sigma &= \left(\frac{(1 - \alpha)c}{Ew^\alpha}\right)^{1/\alpha} \\
&= \left(\frac{(1 - \alpha)c}{Ew^\alpha}\right)^{1/\alpha} \omega_0
\end{align*}
\]

and \( l_0 = \left(\frac{(1 - \alpha)c}{Ew^\alpha}\right)^{1/\alpha} \bar{w} \).

**Definition 3** \( \omega_0 = \bar{w}/l_0 = \left[\frac{c}{(1-\alpha)}\right]^{1/\alpha} \) is the optimal irrigation intensity under certainty. Thus, \( l_\sigma = \frac{(Ew^\alpha)^{1/\alpha}}{\omega_0} \).

Using definitions 2 and 3, surplus is given by

\[
V_S = l_\sigma E [f(w/l_\sigma) - c] = Ew f'(w/l_\sigma) = \alpha \omega_0^{\alpha - 1} (Ew^\alpha)^{1/\alpha}
\]

\(^{16}\)Based on similar considerations, the iso-elastic utility function has been widely used in the precautionary savings literature (see Deaton, 1992).
where the second line follows from $E g(w/l_\sigma) = c$. Self-cultivation is constrained if $l = a$, in which case surplus is

$$V_A = a^{1-\alpha} E w^\alpha - ca.$$ 

### 3.6.2 Leasing

Given the efficiency cost of leasing, area irrigated is lower than it would be under unconstrained self-cultivation.\(^{17}\) In particular, $a + \ell_L = \left(\frac{(1-\alpha)}{c(1+\gamma)} E w^\alpha\right)^{1/\alpha} = (1 + \gamma)^{-1/\alpha} l_\sigma < l_\sigma$. Moreover, surplus is

$$V_L = (1 + \gamma)^{1-1/\alpha} V_S + \gamma ca.$$ 

### 3.6.3 Seasonal contract

The first-order conditions give

$$\alpha E \left(\frac{w - \tau}{a}\right)^{\alpha - 1} = p$$

$$(1 - \alpha) \left(\frac{\tau}{\alpha}\right) = c$$

$$\alpha \left(\frac{\tau}{\alpha}\right)^{\alpha - 1} = p.$$ 

From the second and third equations, we obtain the contract price $p_C = \alpha \omega_0^{\alpha-1}$, which does not depend on $H$, the distribution of $w$. Substituting this expression, along with $\tau = \omega_0 \ell$, into the first condition gives $E (w - \omega_0 \ell)^{\alpha - 1} = (\omega_0 a)^{\alpha - 1}$, the solution to which is $\ell_C = \ell_C(\alpha, a, \omega_0; H)$. The value of the seasonal contract is given by the returns to the well owner; since the $PC$ is binding, the water-seller gets all the surplus. Hence,

$$V_C = a E \left[f \left(\frac{w - \tau_C}{a}\right) - c\right] + p_C \tau_C$$

$$= a^{1-\alpha} E (w - \omega_0 \ell_C)^\alpha - ca + \alpha \omega_0^\alpha \ell_C.$$

\(^{17}\)Recall that cultivation costs $c$ include the (forgone) rental cost of the land even for owned land. Thus, leasing does not entail an additional cost in this respect.
3.6.4 Ex-post bargaining

Two conditions determine allocations in the per-irrigation case: \( \eta E g(\tau/\ell) = c \) and \( f'\left(\frac{w-\tau}{a}\right) = f'\left(\frac{w}{a+\ell}\right) \), the latter implying that \( \tau_P = \frac{\ell}{a+\ell} w \). Solving \( \eta(1-\alpha) E \left( \frac{w}{a+\ell} \right)^\alpha = c \) thus yields \( a + \ell_P = \left( \frac{\eta(1-\alpha)}{c} Ew^\alpha \right)^{1/\alpha} = \eta^{1/\alpha} \ell_\sigma < \ell_\sigma \). This establishes that less area will be cultivated under ex-post bargaining for water as compared to unconstrained self-cultivation. To compute the value of selling on a per-irrigation basis, we consider the joint surplus of both parties to the transaction,

\[
V_P = a E \left[ f\left( \frac{w - \tau_P}{a} \right) - c \right] + \ell_P E \left[ f\left( \frac{\tau_P}{\ell_P} \right) - c \right] \\
= (a + \ell_P) E \left[ f\left( \frac{w}{a + \ell_P} \right) - c \right] \\
= \delta \alpha \omega_0^{\alpha-1} (Ew^\alpha)^{1/\alpha} \\
= \delta V_S
\]

where \( \delta = \frac{1-\eta(1-\alpha)}{\alpha} \eta^{1/\alpha-1} < 1 \). Thus, \( V_P < V_S \), with the relative magnitude of the distortion decreasing (\( \delta \) increasing) in both \( \alpha \) and \( \eta \).

Arguably, the choice of per-irrigation over alternative arrangements would depend on the water seller’s surplus (net value of own production plus revenues from water sales) rather than the joint surplus, \( V_P \). This, however, runs counter to our assumption that all ex-ante negotiations are efficient. In other words, situations in which the per irrigation arrangement yields the highest joint surplus but fails to maximize the well-owner’s surplus would be resolved through side-payments.\(^{18}\)

3.6.5 Comparative statics

We obtain analytical results for the effect of higher uncertainty for most variables of interest:

\textbf{Result 1} \( \partial V_S / \partial \sigma_w^2 < 0 \) and \( \partial \ell_\sigma / \partial \sigma_w^2 < 0 \); likewise for \( V_L, V_P, \ell_L, \) and \( \ell_P \).

\textbf{Proof.} Follows directly from a second-order Taylor expansion of \( Ew^\alpha \) around \( \overline{w} \). \( \blacksquare \)

In the case of the seasonal contract, we have

\textbf{Result 2} \( \partial \ell_C / \partial \sigma_w^2 < 0 \).

\(^{18}\)Of course, lump-sum transfers would not affect marginal choices, such as the extent of land cultivation.
Proof. Follows from a second-order Taylor expansion of \((w - \omega_0 \ell_C)^{\alpha-1}\) around \(\bar{w} - \omega_0 \ell_C\). ■

However, it is not possible to prove that \(\partial V_C / \partial \sigma_w^2 < 0\) for all \(\sigma_w^2\). Instead, we have

**Result 3** If \(\bar{\sigma}_w^2\) is the variance at which \(\ell_C = 0\) (or, equivalently, at which \((E w^{\alpha-1})^{1/(\alpha-1)} = \omega_0 a\)) and if \(\hat{\sigma}_w^2\) is the variance at which \(I_s = a\) (or, equivalently, at which \((E w^\alpha)^{1/\alpha} = \omega_0 a\)), then \(\bar{\sigma}_w^2 < \hat{\sigma}_w^2\).

Proof. Given that \((E w^{\alpha-1})^{1/(\alpha-1)}\) is increasing in \(\sigma_w^2\), whereas \((E w^\alpha)^{1/\alpha}\) is decreasing, it is sufficient to show that \((E w^{\alpha-1})^{1/(\alpha-1)} > (E w^\alpha)^{1/\alpha}\). But, by Jensen’s inequality, \((E w^{\alpha-1})^{1/(\alpha-1)} > \bar{w} > (E w^\alpha)^{1/\alpha}\).

In other words, the seasonal contract, efficient under certainty, at sufficiently high uncertainty becomes no better than autarky and it does so before (i.e., at a lower variance than) unconstrained self-cultivation.

For a more detailed, albeit less general, characterization of the central tradeoff between ex-ante and ex-post inefficiency, we turn to numerical simulations. In particular, we consider a mean-preserving spread based on a 3-point distribution for \(H(w)\), which allows us to vary \(\sigma_w\), holding \(\bar{w}\) constant. The resulting values of \(V_L, V_C, V_P,\) and \(V_A\) for alternative parameter configurations are plotted against \(\sigma_w\) in figure 6. As uncertainty increases, the value of all irrigation arrangements decline and converge to \(V_A\). For sufficiently high variance, transfers of water off the well-owner’s plot by any mechanism become sub-optimal.19

In scenario (1), which assumes \((\alpha, \eta, c) = (0.6, 0.75, 0.5)\), we obtain a crossing of \(V_C\) and \(V_P\); \(V_C\) is maximal at low \(\sigma_w\), whereas \(V_P\) is maximal at high \(\sigma_w\). Greater uncertainty exacerbates the ex-post inefficiency of the seasonal contract, which eventually dominate the ex-ante disincentives under per-irrigation sales.

Such a crossing need not occur, however, as shown in scenario (2) with \((\alpha, \eta, c) = (0.4, 0.75, 0.5)\). In this case, per-irrigation sales are the optimal arrangement for all \(\sigma_w\). Intuitively, as \(\alpha\) falls, the borewell owner expands irrigated area, which, under a seasonal contract, involves committing a larger water transfer to the buyer in all states of the world. As soon as \(\sigma_w\) becomes nonzero, however, the amount of water that can be promised must fall to that which is available in the worst state, \(w_1\). Because of this constraint, \(V_C\) drops down by more than \(V_P\).

Scenario (3), in which \((\alpha, \eta, c) = (0.6, 0.6, 0.5)\), shows that lowering the bargaining power of the water-buyer reduces the cultivation incentive under the per-irrigation arrangement.

---

19Whether this convergence point is reached before \(\sigma_w\) achieves its maximum depends on the parameters of the model, plot area, and \(H\).
If there is still a crossing between $V_C$ and $V_P$ it occurs at a much higher $\sigma_w$ and possibly not before water sales of any kind become sub-optimal. Finally, in scenario (4), we lower the costs of cultivation relative to scenario (3) so that $(\alpha, \eta, c) = (0.6, 0.6, 0.4)$. Here, all the irrigation arrangements yield higher surplus, but $V_L$ and $V_P$ in particular rise relative to $V_C$ and, especially, relative to $V_A$.

4 Estimation framework

Table 2 enumerates a well-owner’s choices in terms of the notation developed in the previous section (see cols. 1 and 2). The last two columns of the table provide descriptive statistics for the sample of 2423 borewells. Roughly half of the borewell owners transferred groundwater in some form during the past *rabi* season, most often selling it under a seasonal contract, followed by per-irrigation sales. Lagging far behind, in terms of frequency, are groundwater transfers to leased plots within the adjacency. Conditional on making a transfer, mean area irrigated (aside from that of the well-owner’s plot) is comparable across seasonal contract and leasing, but substantially lower for the per-irrigation arrangement.

4.1 Additional parameterizations

4.1.1 Fixed transactions costs

To explain why a well-owner would irrigate *exactly* his own plot area (i.e., $l = a$) instead of making a water transfer, however minute, to another plot, we need to introduce a fixed transactions cost, $\kappa$. One can think of $\kappa$ as reflecting in part the availability of adjacent land to irrigate. Clearly, if all neighboring plots already have their own borewell, then arranging a remunerative lease or water sale would be a challenge. However, the presence of a borewell on a neighboring plot may itself be driven by the failure of groundwater markets. In other words, in areas less conducive to such markets, there may be a greater incentive for farmers to drill wells of their own. To capture heterogeneity in $\kappa$ across adjacencies, therefore, we also need, in effect, an instrument for borewell density.

Thankfully, as figure 5 suggests, we do have such an instrument: average plot size in the adjacency, $\bar{a}_i$, which is plausibly unrelated to groundwater markets except through its strong

---

20 As noted earlier, for the most part there is a one-to-one correspondence between borewells and adjacencies, but we do have more than a hundred cases of multiple wells in the reference plot. In this situation, we allocate the total area of the adjacency equally among wells, treating each well as an independent decision unit within its own (pro-rated) adjacency.
positive association with borewell density. Thus, we incorporate heterogeneity in irrigable land availability by introducing $\pi_i$ into the structural model through $\kappa$. Finally, we allow for the possibility that arranging a lease may be more costly than arranging a water sale for the simple reason that the former likely involves not only finding an adjacent plot without its own irrigation but also finding a plot whose owner has income-generating opportunities outside of agriculture. Putting these two elements together, we have

$$\kappa_{ij} = \kappa_j \exp(\beta \pi_i)$$

for water transfer type $j$, where $\kappa_L \neq \kappa_C = \kappa_P = \kappa_T$. In particular, we expect that $\kappa_L > \kappa_T$.

### 4.1.2 Water flow

Well-owners report probabilities for five water flow states, corresponding to “full”, $3/4$, $1/2$, $1/4$, and zero flow. Since water discharge is proportional to the square of pipe radius $r_i$, we have

$$w_{ki} = \lambda r_i^k$$

for $k = 0, 1, 2, 3, 4$, (3)

where $\lambda$ is a parameter. It is straightforward to verify that, like the scale of the production function ($\zeta$), $\lambda$ is not identifiable; hence, we normalize it to $\frac{1}{4}$.

### 4.1.3 Cost disturbance

To explain why different water transfer arrangements (including no transfers at all) are chosen across observationally equivalent borewells/adjacencies, as well as why different areas are cultivated conditional on the transfer arrangement, we introduce a cost disturbance $\varepsilon$ such that $c = \bar{c} e^{\varepsilon}$. To be clear, cost heterogeneity is assumed to reflect variation in local (adjacency-level) conditions, such as soil texture, depth, and water-retention capacity, rather than characteristics of the *cultivator*.

### 4.2 Likelihood function

Our model delivers choice-specific residuals $\varepsilon_{ij}$ for $j = S, L, C, P$ such that $\ell_{ij} = \ell_j(\varepsilon_{ij})$, where $\ell_{ij}$ is the observed area under that arrangement for borewell $i$ (see Table 2). In

\footnote{As noted earlier, our model of groundwater transactions assumes that cultivators share a common production and cost function.}
addition, we have thresholds,

\[
V_A(\varepsilon_{jA}) = V_j(\varepsilon_{jA}) - 1(j \neq S)\kappa_{ij} \quad j = S, L, C, P
\]
\[
V_j(\varepsilon_{jk}) - \kappa_{ij} = V_k(\varepsilon_{jk}) - \kappa_{ik} \quad (j, k) = \{(L, C), (L, P), (P, C)\},
\]
defined by the crossings of the value functions on the support of \(\varepsilon\). These residuals and thresholds embed the entire content of the theory. To illustrate the role of the latter, Figure 7 shows one of the 24 mutually exclusive and exhaustive partitions of the support of \(e^\varepsilon\) into regions where either per-irrigation, leasing, seasonal contracts, autarky, or unconstrained self-cultivation uniquely provides the greatest surplus.

Let \(Pr(j|\Theta, Z_i)\) denote the probability of choice \(j\) conditional on the parameters of the model \(\Theta = (\alpha, \eta, \gamma, \bar{c}, \bar{r}_L, \bar{r}_T, \beta)\) and the data \(Z_i = (\pi_{ki}, r_i^2, a_i, \bar{a}_i)\). If we assume that \(\varepsilon_i/\sigma\) is i.i.d. standard normal with c.d.f. \(\Phi\), then

\[
Pr(S|\Theta, Z_i) = 1 - \Phi(\varepsilon_{iSA}/\sigma)
\]
\[
Pr(A|\Theta, Z_i) = \Phi(\varepsilon_{iSA}/\sigma) - \Phi(\varepsilon_{iA}/\sigma), \quad \text{where} \quad \varepsilon_{iA} = \max\{\varepsilon_{iLA}, \varepsilon_{iCA}, \varepsilon_{iPA}\}
\]
\[
Pr(j|\Theta, Z_i) = \sum_r \left[\Phi(\varepsilon_{ijr}^{(1)}/\sigma) - \Phi(\varepsilon_{ijr}^{(1)}/\sigma) + \Phi(\varepsilon_{ijr}^{(2)}/\sigma) - \Phi(\varepsilon_{ijr}^{(2)}/\sigma)\right] D_{ir} \quad j = L, C, P
\]

where the \(\varepsilon_{ijr}^{(m)}\) and \(\varepsilon_{ijr}^{(m)}\) are, respectively, upper and lower limits of the regions of integration for the probability of choice \(j\) in partition regime \(r\), and \(D_{ir} = 1\) if observation \(i\) lies in regime \(r\), zero otherwise. Appendix Table A.1 provides the correspondence between each of these limits and the value function crossings \(\varepsilon_{jA}\) and \(\varepsilon_{jk}\). Note that for \((j, k) = \{(L, C), (P, C)\}\) there can be two value function crossings, enumerated here by the index \(m\). Also, for well-owners who choose constrained own-cultivation \((j = A)\), we do not know the transfer arrangement they would have made. Hence, the lower limit of integration for their choice probability, \(\varepsilon_{iA}\), represents the highest \(\varepsilon_i\) that would have induced them to make a water transfer of any type.

Now, letting \(d_{ij}\) be an indicator that takes a value of 1 when well-owner \(i\) chooses water arrangement \(j = S, A, L, C, P\), the log-likelihood is

\[22\text{In some regimes, } j \text{ may not be optimal for any } \varepsilon_i, \text{ in which cases we penalize the likelihood by setting the choice probability to a very small number.}\]
\[ \mathcal{L}(\Theta) = \sum_i \left\{ \sum_{j \neq A} d_{ij} \log \left[ \Pr(j|\Theta, Z_i) \phi(\varepsilon_{ij}/\sigma) / \sigma \right] + d_{iA} \log \left[ \Pr(A|\Theta, Z_i) \right] \right\}, \]

where \( \phi \) is the standard normal density.

### 4.3 Identification

To gain intuition for identification of the structural parameters, consider just the subsample of well-owners engaged in unconstrained self-cultivation (i.e., \( d_{iS} = 1 \)). Taking logs of equation (2) and using (3), we have, for \( \varepsilon_i > \varepsilon^{*}_{SA} \),

\[
\log \left[ \frac{(a_i - \ell_{iS})}{r_i^2} \right] = \frac{1}{\alpha} \left[ \log(1 - \alpha) - \log \bar{c} + \log \sum_k \pi_{ki} \left( \frac{k}{4} \right)^\alpha - \varepsilon_i \right].
\]

Thus, \( \alpha \) controls how fast irrigated area (scaled by well capacity) falls with variability in water supply, \( \bar{c} \) can be extracted from the intercept, and \( \sigma^2 \) is the residual variance.

Of course, \( \alpha \) is identified off of other moments of the data as well; in particular, the shares of well-owners choosing different water arrangements conditional on their water supply variability in addition to the average area irrigated under each arrangement. Similarly, the water-buyer’s bargaining weight \( \eta \) and the leasing cost \( \gamma \) are identified off of the relative shares of water transactions done on a per-irrigation or leasing basis and by differences in average land area irrigated across transaction types. Finally, the transaction cost parameters \( \kappa_L, \kappa_T, \) and \( \beta \) are identified off the proportion of well-owners who irrigate their full plot yet choose not to transfer any water, how this proportion varies with average plot size in the adjacency, as well as off of the relative proportions of leasers and water-sellers.

### 5 Results

#### 5.1 Parameter estimates

Parameter estimates for a restricted model, with \( \beta = 0 \), are reported in Table 3 along with standard errors.\(^{24}\) The estimates appear reasonable for the most part and are precise. In particular, \( \alpha \) is considerable less than one, whereas a value close to one would have implied

\[ \varepsilon^{*}_{SA} = \frac{1}{\alpha} \left[ \log(1 - \alpha) - \log \bar{c} + \log \sum_k \pi_{ki} \left( \frac{k}{4} \right)^\alpha \right] + 2 \log r_i - \log(a_i). \]

\(^{23}\)Note that \( \varepsilon^{*}_{SA} = \frac{1}{\alpha} \left[ \log(1 - \alpha) - \log \bar{c} + \log \sum_k \pi_{ki} \left( \frac{k}{4} \right)^\alpha \right] + 2 \log r_i - \log(a_i). \)

\(^{24}\)Given the computation time involved in maximizing the likelihood, we drew a random 25% sample, stratified by choice for our preliminary runs. In subsequent versions of the paper, we will present full sample results.
little role for groundwater uncertainty. The marginal inefficiency of leasing land, \( \gamma \), is modest at around 5% of cultivation cost. Our estimate of \( \eta \) suggests that water-buyers’ bargaining share is around three-quarters. Finally, we find as expected that \( \kappa_L > \kappa_T \).

5.2 Within sample fit

To assess how well the model fits observed choices, we predict for each borewell owner a probability for each \( j = S, A, L, C, P \). Table 4 provides sample means of these predictions by actual choice. The model does quite a good job predicting borewell owners’ choices of \( P \) and \( C \), but less well on the other choices.

5.3 Counterfactuals

Next, we consider a counterfactual in which borewell owners are endowed with plots of essentially unlimited size. In other words, we ask what would happen in a world where landholdings were no longer fragmented. In terms of our model, land consolidation is tantamount to having each borewell owner earn \( V_S \) rather than whichever \( V_j \) is maximal given his actual plot size, pipe-width, and model parameters.

In figure 8, we plot the change in both irrigated area and surplus or profits after the hypothetical consolidation as a function of (log of) plot area adjusted for pipe-width (i.e., \( a_i/r_i^2 \)). The gains to consolidation are decreasing in adjusted plot size, since borewell owners with larger plots, of course, need resort less to costly groundwater transactions.

6 Conclusion

We have developed a model of groundwater transactions in the spirit of Hart and Moore’s (2008) contracts as reference-points theory, which emphasizes the role of payoff uncertainty, in our context arising from unpredictable fluctuations in groundwater supply. Our structural estimates of the model using micro-data on area irrigated under different transaction types are reasonable and allow for interesting counterfactual exercises. In particular, we use the estimated parameters to quantify the potential gains from land consolidation, which are found to be substantial in terms of irrigated area, if not surplus, and highest for borewell owners with the smallest plots.
References


Figure 1: Borewell-level groundwater uncertainty by District
Figure 2: Presence of a borewell and area of plot

Figure 3: Presence of a borewell and area of plot controlling for wealth
Figure 4: Dry-season fallow and area of plot
Figure 5: **Borewell density, groundwater transfers, and average plot size**
Figure 6: Model simulations under alternative parameter configurations

Common parameters: gamma=0.15; a=1; w1=1; w2=4; w3=6
Figure 7: A possible partition of the error-space
Figure 8: IMPACT OF HYPOTHETICAL LAND CONSOLIDATION
### Table 1: Precautionary Planting and Risk Aversion

<table>
<thead>
<tr>
<th></th>
<th>Subjective measure ($Y^s$)</th>
<th>Objective measure ($Y^o$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td><strong>(a) Baseline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log($CV$)</td>
<td>(-0.0849^{***})</td>
<td>(-0.0678^{***})</td>
</tr>
<tr>
<td></td>
<td>(0.0151)</td>
<td>(0.0157)</td>
</tr>
<tr>
<td>log(pipe width)</td>
<td>(-0.0657^{***})</td>
<td>(-0.0543^{***})</td>
</tr>
<tr>
<td></td>
<td>(0.0178)</td>
<td>(0.0186)</td>
</tr>
<tr>
<td><strong>(b) $RISK^1$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log($CV$)</td>
<td>(-0.0997^{**})</td>
<td>(-0.0842^{**})</td>
</tr>
<tr>
<td></td>
<td>(0.0413)</td>
<td>(0.0415)</td>
</tr>
<tr>
<td>log(pipe width)</td>
<td>(-0.0656^{***})</td>
<td>(-0.0540^{***})</td>
</tr>
<tr>
<td></td>
<td>(0.0179)</td>
<td>(0.0187)</td>
</tr>
<tr>
<td>$RISK^1$</td>
<td>0.00390</td>
<td>0.00408</td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0129)</td>
</tr>
<tr>
<td>log($CV$)×$RISK^1$</td>
<td>0.00243</td>
<td>0.00268</td>
</tr>
<tr>
<td></td>
<td>(0.00657)</td>
<td>(0.00656)</td>
</tr>
<tr>
<td><strong>(c) $RISK^2$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log($CV$)</td>
<td>-0.0543</td>
<td>-0.0301</td>
</tr>
<tr>
<td></td>
<td>(0.0367)</td>
<td>(0.0372)</td>
</tr>
<tr>
<td>log(pipe width)</td>
<td>(-0.0664^{***})</td>
<td>(-0.0552^{***})</td>
</tr>
<tr>
<td></td>
<td>(0.0178)</td>
<td>(0.0186)</td>
</tr>
<tr>
<td>$RISK^2$</td>
<td>-0.0955</td>
<td>-0.117</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.107)</td>
</tr>
<tr>
<td>log($CV$)×$RISK^2$</td>
<td>(-0.0471)</td>
<td>(-0.0580)</td>
</tr>
<tr>
<td></td>
<td>(0.0550)</td>
<td>(0.0551)</td>
</tr>
</tbody>
</table>

| Controls | No | Yes | Yes | No | Yes | Yes |
| District dummies | No | No | Yes | No | No | Yes |
| Observations | 2,396 | 2,388 | 2,388 | 2,417 | 2,409 | 2,409 |

**Notes:** Robust standard errors in parentheses adjusted for clustering on borewell (*** p<0.01, ** p<0.05, * p<0.1). CV is the coefficient of variation of end-of-season well flow; $RISK^1$ is a self-assessed ranking of risk tolerance; $RISK^2$ is an index of marginal willingness-to-pay for risk based on a Binswanger lottery. Unreported controls are as follows: pump horse-power, log of well depth, number of other borewells within 100 meters, dummy for presence of recharge source.
Table 2: Choices, Notation, and Sample Characteristics

<table>
<thead>
<tr>
<th>Choice</th>
<th>Model Notation</th>
<th>Descriptive Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>Area irrigated</td>
</tr>
<tr>
<td>S: Unconstrained self-cultivation</td>
<td>$V_S$</td>
<td>$a - \ell_S$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: Constrained self-cultivation</td>
<td>$V_A$</td>
<td>$a$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L: Leasing</td>
<td>$V_L - \kappa_L$</td>
<td>$a + \ell_L$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C: Seasonal contract</td>
<td>$V_C - \kappa_T$</td>
<td>$a + \ell_C$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P: Per-irrigation sale</td>
<td>$V_P - \kappa_T$</td>
<td>$a + \ell_P$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Sample means (standard deviations) in columns 4 and 5.
Table 3: Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coef.</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.710</td>
<td>0.0026</td>
</tr>
<tr>
<td>$\bar{c}$</td>
<td>0.233</td>
<td>0.0093</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.045</td>
<td>0.0027</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.765</td>
<td>0.0067</td>
</tr>
<tr>
<td>$\kappa_L$</td>
<td>0.100</td>
<td>0.0103</td>
</tr>
<tr>
<td>$\kappa_T$</td>
<td>0.096</td>
<td>0.0098</td>
</tr>
<tr>
<td>$\sigma_\varepsilon$</td>
<td>0.833</td>
<td>0.0271</td>
</tr>
</tbody>
</table>

Mean log-likelihood 2.358

Notes: 25% sample with $\beta = 0$.

Table 4: Within Sample Prediction

<table>
<thead>
<tr>
<th>Choice</th>
<th>S</th>
<th>A</th>
<th>L</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>S: Unconstrained self-cultivation</td>
<td>64.97</td>
<td>13.18</td>
<td>0.19</td>
<td>10.29</td>
<td>11.35</td>
</tr>
<tr>
<td>A: Constrained self-cultivation</td>
<td>40.33</td>
<td>15.08</td>
<td>1.60</td>
<td>17.55</td>
<td>25.06</td>
</tr>
<tr>
<td>L: Leasing</td>
<td>34.02</td>
<td>13.05</td>
<td>1.39</td>
<td>17.91</td>
<td>33.63</td>
</tr>
<tr>
<td>P: Per-irrigation sale</td>
<td>27.74</td>
<td>12.61</td>
<td>3.12</td>
<td>22.15</td>
<td>34.37</td>
</tr>
<tr>
<td>C: Seasonal contract</td>
<td>26.48</td>
<td>9.91</td>
<td>2.54</td>
<td>23.64</td>
<td>37.43</td>
</tr>
<tr>
<td>Data</td>
<td>26.90</td>
<td>26.90</td>
<td>4.62</td>
<td>15.02</td>
<td>26.57</td>
</tr>
<tr>
<td>Model</td>
<td>41.09</td>
<td>12.73</td>
<td>1.69</td>
<td>17.92</td>
<td>26.45</td>
</tr>
</tbody>
</table>

Notes: Figures are percentages. Data refers to raw sample percentage by choice and Model refers to average predicted choice probabilities unconditional on actual choice.
Appendix

A Additional Figures and Tables
Figure A.1: Rising Importance of Borewell Irrigation in India

Source: ICRISAT district level database.
(All India covers 17 major states, excluding W. Bengal and Maharashtra for lack of data)
Figure A.2: Average depth to water table in Andhra Pradesh
Figure A.3: Increasing Land Fragmentation in India