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

Xavier Giné and Hanan G. Jacoby*

November 13, 2014

Abstract

Access to groundwater has been a key driver of agricultural productivity growth and rural poverty reduction in South Asia. Yet, markets for groundwater have not developed everywhere. We develop a contract-theoretical model of groundwater transactions under payoff uncertainty, which arises from unpredictable fluctuations in groundwater availability during the agricultural 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 borewell discharge collected from over 1,600 well-owners across four districts in southern India. We use the estimates to quantify the contracting distortion and its impact on the development of groundwater markets.

*Development Research Group, The World Bank, 1818 H St. NW, Washington DC, 20433. Giné: xgine@worldbank.org; Jacoby: hjacoby@worldbank.org. We owe a particular debt to K.P.C. Rao for his efforts in managing the field work associated with this study, and to his survey team. We also thank Mark Gersovitz for extraordinarily useful comments on an earlier draft, as well as seminar participants and several discussants for their ideas and suggestions. Financial support for this project was provided by the Research Committee of the World Bank. Views expressed in this paper are those of the authors, and do not necessarily reflect the opinions of the World Bank, its executive directors, or the countries they represent.
1 Introduction

A central theme in the economics of organization is that long-term contracts protect investments specific to a trading relationship. The early transactions costs literature (Williamson, 1971; Klein et al., 1978) recognized not only these benefits of long-term contracts but also the costs; in an uncertain environment, contractual rigidity inevitably leads to resource misallocation, a distortion obviated by ex-post or spot contracting. Empirical research within this tradition thus asked two complementary questions: Does the duration of long-term contracts depend on the degree of asset specificity (Joskow, 1987)? And, are long-term contracts structured so as to adapt to changing economic conditions (e.g., Goldberg and Ericson, 1987; Masten and Crocker, 1985; Crocker and Masten, 1988)? But, the choice of long-term over spot contract, and whether it too is driven by this fundamental tradeoff between ex-ante and ex-post inefficiency, has not been rigorously investigated.\(^1\)

To fill this lacuna, we consider an environment–agricultural production in southern India–characterized by substantial upfront investment as well as by uncertainty regarding a critical input into subsequent production stages: groundwater extracted from private borewells, the exclusive source of dry-season irrigation. The link between access to groundwater, or lack thereof, and rural poverty in South Asia (see Shah, 2007, and especially Sekhri, 2013) highlights a growing economics literature on the industrial organization of groundwater.\(^2\) In principle, groundwater markets should serve the irrigation needs of poor farmers who cannot afford their own borewells. However, since water is extremely costly to transport, these markets tend to be highly spatially fragmented and, hence, inherently uncompetitive (Jacoby et al. 2004). In our setting, bilateral transactions between well-owners and neighboring farmers take one of two forms: spot contracts, in which groundwater is sold on a per-irrigation basis, and long-term (i.e., seasonal) contracts, which specify price and area irrigated over an entire crop cycle ex-ante.

We develop a model in which spot contracts are fully state contingent and thus ex-post efficient, but, due to the classic holdup problem, ex-ante inefficient. In particular, planting incentives of water-buyers are distorted. Long-term contracts, by contrast, are assumed

\(^1\)Lafontaine and Slade (2012) provide a thorough review of empirical research on inter-firm contracting from various theoretical perspectives. A strand of the literature does consider the choice between long-term contracts and spot markets (e.g., Carlton, 1979; Polinsky, 1987; Hubbard and Weiner, 1992). However, in these models, there is no relationship-specific investment; firms incur the transactions costs of long-term contracts to insure against cash-flow variability.

to deter holdup, but lead to ex-post inefficiency insofar as groundwater is misallocated across farms once the state of nature is revealed. Our model yields the sharp prediction that as groundwater supply uncertainty increases, long-term contracts become unattractive relative to spot arrangements. In addition, higher uncertainty reduces the overall extent of groundwater markets.

A key contribution of this paper lies in quantifying these contracting distortions, as well as their impacts, using a structural econometric model. A rather unique feature of a groundwater economy that allows us to do this is that buyers and sellers are both agricultural producers, cultivating side-by-side with the same technology, the total returns to which are (to a first approximation) proportional to area irrigated. Gagnepain et al. (2013) is another rare example of structural estimation and quantitative welfare analysis within the empirical contracts literature. While their analysis of the tradeoff between ex-post renegotiation and ex-ante incentives in the context of French public-sector contracts bears a superficial resemblance to ours, contractual choice is driven in their case by asymmetric information rather than, in our case, by what Hart (2009) terms payoff uncertainty. Moreover, in the setting we consider, agents have the option not to contract or trade at all (and many do not), which allows us to investigate how payoff uncertainty affects overall market activity.

Our model of agricultural production under stochastic groundwater supply accounts for the choice between seasonal contracts and per-irrigation sales, for water transfers through leasing, and (crucially) for the areas irrigated under each such arrangement. We have data collected from a large sample of borewell owners across six districts of Andhra Pradesh and Telangana states in southern India. Our specially-designed groundwater markets survey takes particular care to elicit from each respondent a subjective probability distribution of their borewell’s discharge near the end of the season conditional on its initial discharge. The structural parameters of the model are identified principally off of variation across borewells in this conditional probability distribution.

With structural parameter estimates in hand, we perform two distinct model valida-

---

3On this point, we appeal to the recent insight of Hart and Moore (2008) that contracts act as reference points, establishing what each party in the transaction is entitled to. Opportunism thus leads to deadweight losses. In the earlier property rights theory of the firm associated with Grossman and Hart (1986) and Hart and Moore (1990), renegotiation is efficient so that holdup is virtually inevitable (Hart, 1995). As a result, there is no functional difference between contracts agreed upon ex-ante and those agreed upon ex-post. Fehr et al. (2011) and Hoppe and Schmitz (2011) corroborate the reference-point idea experimentally.

4While several recent papers incorporate subjective probabilities into structural econometric models (see Attanasio, 2009, for a literature review and, more recently, 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.
tion exercises, each on a separate sub-sample of the data. First, we retain a *nonrandom* “holdout” sample consisting of two high-uncertainty districts. Borewells from the remaining four districts constitute the “estimation” sample, which we, in turn, split into two *random* sub-samples. On one of these, the “model development” sub-sample, we estimate our econometric model. On the other, the “model validation” sub-sample, we assess model fit. Out-of-sample goodness-of-fit tests are robust to data-driven model development, otherwise known as over-fitting or, more pejoratively, as data-mining. However, as noted by Keane and Wolpin (2007), model predictions based on the validation sub-sample must perform remain within the support of policy regime extant in the estimation sample. Therefore, good out-of-sample fit does not necessarily mean that the model can accurately predict large shifts in policy. In this sense, internal validity does not imply external validity. To assess the latter, we use the nonrandom holdout sample. Keane and Wolpin (2007) argue for choosing “a [holdout] sample that differs significantly from the estimation sample along the policy dimension that the model is meant to forecast (p. 1352).” The analogue, in our setting, to a policy regime (i.e., contracting environment) “well outside the support of the data” is the large difference in levels of groundwater supply uncertainty between borewells in the four estimation sample districts (low uncertainty) as compared to the two holdout sample districts (high uncertainty).

Structural estimation provides a twofold benefit: First, it enables us to assess the performance of actual contractual arrangements against the first-best benchmark. Second, it allows these contracting distortions to be compared quantitatively to other sources of groundwater market failure. We consider, in particular, the problem of *strategic substitutability* among borewells; i.e., once neighboring landowners each have their own well, there is limited scope for trading groundwater between them. In other words, there may be coordination failure, in which case a social planner would optimally choose to drill fewer wells (thus economizing on fixed costs) and require more water-sharing between neighbors. Although it is impossibly complex to integrate the initial well-drilling game, playing out as it does over both time and space, we do incorporate a fixed cost of arranging groundwater transactions into our structural model. We find that this fixed cost, which captures the extent of coordination failure, is in fact quite large relative to the uncertainty-induced contracting distortion.

---

5To our knowledge, this two-level approach to model validation has not been attempted before in the context of structural estimation.

6Foster and Sekhri (2008) argue that groundwater markets might crowd out well-drilling, whereas our concern is with how well-drilling might crowd out groundwater markets. We deal with the potential simultaneity between well-drilling and groundwater markets in Section 4.
The next section of the paper lays out the formal theoretical arguments. Section 3 describes our survey data and the groundwater economy of southern India in greater detail. Section 4 adapts the theoretical model for the purposes of structural estimation and derives the likelihood function. Estimation results and counterfactual simulations are reported in Section 5. Section 6 concludes.

2 Theory

2.1 Preliminaries

We begin by briefly enumerating our assumptions, leaving the more extended justifications for section 3.

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

This presupposes that the outright purchase of neighboring plots is generally infeasible, perhaps due to mortgage finance constraints. At any rate, (A.1) is backed by data, since, as we will see, a large majority of plots are inherited.

A.2) Borewells and groundwater: A reference plot has a borewell drawing a stochastic quantity of groundwater \( w \) over the growing season, where \( w \) has p.d.f. \( \psi(w) \) on support \([w_L, w_H]\). Groundwater is the sole irrigation source.

Property rights to groundwater are not clearly delineated in India, so there is no legal limit to withdrawals. Because electricity is provided free at the margin, farmers run their pumps for the maximum number of hours that power is available on any given day. Aside from this constraint, \( w \) depends on the capacity of the well (pipe-width), availability of groundwater in the aquifer, and the local hydrogeology.

A.3) Technology: The common crop output production function, \( y = F(l, w, x) \), depends on three inputs: land \( l \), seed \( x \), and water \( w \), with land and seed used in fixed proportions. For any level of \( x \), \( y/l = f(w/l) \equiv f(\omega) \), where \( \omega \) is irrigation intensity and the intensive production function, \( f \), is increasing, concave, with \( f(0) = 0 \).

\( ^7 \)Constant returns to scale is both technically convenient and empirically sensible. 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.
Given (A.3), we may write net revenue as \( l\{f(\omega) - c\} \), where \( c \) is the cost of the required seed per acre cultivated.

**A.4) Risk preferences:** Farmers are risk neutral.

We defer a discussion of the role of risk preferences to the next section, only noting here that risk-neutrality is a core assumption of the transactions cost literature.

**A.5) Adjacent land:** A well-owner is not limited in the area of adjacent land that his borewell can irrigate.

In invoking (A.5), we abstract from any demand-side constraints such as may arise when most or all adjacent landowners also have their own borewells. While this assumption vastly simplifies the theoretical analysis, it is is clearly unrealistic and, for this reason, it will be relaxed in the empirical implementation.

Consider, first, a well-owner’s choice of area cultivated (irrigated) when his own plot size is not a constraint. Let \( \ell_U = \arg \max \{ Ef(w/l) - c \} \) and define the marginal return as

**Definition 1** \( g(\omega) = f(\omega) - \omega f'(\omega) \).

The necessary condition for optimal planting

\[
Eg(\omega) = c
\]  

(1)
equates the expected marginal return to the marginal cost of cultivation.

Now, letting \( r \) index mean preserving increases in groundwater supply uncertainty, we have

**Proposition 1 (Precautionary planting)** If \( g \) is strictly concave, then \( \partial \ell_U / \partial r < 0.8 \).

In other words, 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 increases in supply uncertainty by committing less area to irrigate.

The surplus generated by a borewell under unconstrained self-cultivation is

**Definition 2** \( V_U = \ell_U E [f(w/\ell_U) - c] \).

---

8*Proof:* Follows directly from Theorem 1 of Diamond and Stiglitz (1974).
In case $\ell_U > a$, we may think of $V_U$ as the surplus derived by the well-owner if he could sell an unlimited amount of groundwater in a competitive spot market. As mentioned, however, groundwater transactions do not resemble this competitive, arm’s-length, ideal.

We now consider the two observed forms of bilateral contracting between well-owner and buyer using Table 1 to guide and unify the discussions in the next two subsections.

### 2.2 Long-term contracts

The canonical long-term contract 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), we think of such (ex-ante) contracts as establishing entitlements. Ex-post renegotiation of the terms, or hold-up, will therefore lead to deadweight losses due to aggrievement by one or both parties. To bring the tradeoff between ex-ante and ex-post inefficiency into stark relief, we assume that these deadweight losses make hold-up prohibitively costly. 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.

Let $\tau$ denote the total transfer of groundwater at per unit price $p$ to irrigate a field of size $l$. The optimal simple (i.e. single-price) contract solves

\[
\max_{p,l,a} \left\{ Ef \left( \frac{w - \tau}{a} \right) - c \right\} + p\tau \quad \text{s.t.} \quad PC : l \left\{ f \left( \frac{\tau}{l} \right) - c \right\} - p\tau \geq 0
\]

\[
IC : \tau = \arg \max_{\tau \in [0,\omega_L]} l \left\{ f \left( \frac{\tau}{l} \right) - c \right\} - p\tau
\]

---

9 To see why, let subscripts $b$ and $s$ denote water-buyer and seller, respectively. Further, let $p$ be the spot price and $\ell_b$ the buyer’s cultivated area such that $\ell_U = a + \ell_b$. It is easy to see that $f'(\omega_b) = f'(\omega_s) = p$ which implies that $\omega_b = \omega_s$. Thus, $V_U = E \left[ a(f(\omega_s) - c) + p\omega_b\ell_b \right] = E \left[ a(f(\omega_s) - c) + f'(\omega_b)\omega_b\ell_b \right] = E \left[ a(f(\omega_s) - c) + \ell_b(f(\omega_b) - c) \right] = E \left[ (a + \ell_b)(f(\omega_s) - c) \right]$, where the penultimate expression follows from $Eg(\omega_b) = c$, the necessary condition for the buyer’s optimal area cultivated.

10 More precisely, this literature assumes 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). Alternatively, we can follow Gagnepain et al. (2013) and appeal to relational contracting to explain the existence of nontrivial long-term contracts.
The first term in the well-owner’s objective function (top line) is the expected revenue from crop production on his own plot net of cultivation costs, which is diminished when he sells water to a neighbor; the second term is his total revenue from the sale. The participation constraint (PC) stipulates that the crop revenue of the buyer net of both cultivation and water costs cannot be negative. Finally, the incentive constraint (IC) says that the transfer is maximizing the buyer’s net revenue, subject to the constraint that the promised amount cannot exceed the available supply of water in the lowest state of the world, $w_L$. Note that expectations are dropped in both the PC and IC because, under the contract, $l$ and $\tau$ are fixed ex-ante. Thus, the seasonal contract offers an assured supply of irrigation to the buyer; the direct cost of production variability is borne fully by the seller on his plot.

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

\[
Ef'\left(\frac{w - \tau}{a}\right) = p \\
g\left(\frac{\tau}{l}\right) = c \\
f'\left(\frac{\tau}{l}\right) = p, \tag{3}
\]

the solution to which is the water transfer-area pair $(\tau_C, l_C)$. Divergence of supply and demand for irrigation ex-post creates a distortion. Since (3) implies $Ef'\left(\frac{w - \tau}{a}\right) = f'\left(\frac{\tau_C}{l_C}\right)$, it is not true, in general, that $f'\left(\frac{w - \tau}{a}\right) = f'\left(\frac{\tau_C}{l_C}\right) \forall w$, which would obtain if $\tau$ were state-contingent, as in a competitive spot market (see fn. 9). It follows as a corollary that the distortion vanishes as $\psi$ becomes degenerate, in which case $g\left(\frac{l_C}{\tau_C}\right) = g\left(\frac{w}{l_C+a}\right) = c = g\left(\frac{w}{l_U}\right)$ which implies that $l_C = l_U - a$. Thus, in the absence of uncertainty, the amount of land irrigated and the economic surplus generated by the borewell would be the same under the seasonal contract as under a competitive spot market; i.e., the long-term contract would achieve the first-best.

---

11 Without loss of generality, we assume that the constraint that the water-seller’s cultivated area $l_s$ cannot exceed his plot area $a$ is binding; i.e., the well-owner always fully cultivates his land before selling any groundwater. Proof: Suppose not, then the optimal choice of $l_s$ requires $Eg\left(\frac{w - \tau}{l_s}\right) = c$. However, equation (1) $\Rightarrow Eg\left(\frac{w}{l_C}\right) = c \Rightarrow \tau = w(1 - l_s/l_U)$, which is a contradiction because $\tau$ cannot be state-contingent.
2.3 Spot contracts

Groundwater may also be sold on a per-irrigation basis. Once the season is underway, however, commitments have been made. The potential seller has retained (i.e., refrained from contracting out) the rights to some excess water from his well during the season whereas the potential buyer has planted a crop in an adjacent plot.\textsuperscript{12} Since each party has some degree of bargaining power, we use a Nash bargaining framework.

To be clear, in a per-irrigation arrangement, there is 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.\textsuperscript{13}

Returning to the ex-post stage, let $\tilde{\tau}$ 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 $u = lf((\tilde{\tau} + \Delta)/l) - p\Delta$, whereas that of the seller is $v = af((w - \tilde{\tau} - \Delta)/a) + p\Delta$. The no-trade payoffs are given by $\underline{u} = lf(\tilde{\tau}/l)$ and $\underline{v} = af((w - \tilde{\tau})/a)$, respectively. The absence of $c$ in these payoff functions reflects the fact that all cultivation costs have already been incurred.

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

\textsuperscript{12}An early descriptive study of groundwater markets in a Tamil Nadu village captures the buyer’s predicament: “A particularly potent source of control which a water seller can exercise over a water purchaser is that the former is in a position to stop supplying water to the latter at a crucial stage of crop growth, on alleged grounds of his pumpset being in a broken down state, or some such similar, transparently flimsy excuse. A cessation of the provision of water after an initial supply lasting for a month or forty-five days could expose the water purchaser, who by then would have invested substantially in raising the crop, to the danger of heavy losses. Neither is the purchaser free to turn to an alternative supplier of water: a customary practice in force in the village ensures that a water seller may supply water only to owners of contiguous plots...” (Janakarajan, 1993, p. 71)

\textsuperscript{13}We also abstract from any reallocation of property rights between the parties at this stage that redistributes ex-post bargaining power à la Grossman and Hart (1986). Later, in the empirical model, we allow for one form of vertical integration: the well-owner can lease an adjacent plot without a well of its own, although this too is costly.
\[ \eta \left[ \frac{f(w - \tau - \Delta_a)}{a} - f(w - \tau_a) \right] - (1 - \eta) \left[ \frac{f(\tau + \Delta_l)}{l} - f(\tau_l) \right] + p = 0 \]
2.4 Other contracts

Our approach, following in the tradition of the empirical contracts literature (e.g., Gagnepain et al., 2013), has been to model only the principal arrangements observed in the data. Nevertheless, it is worth a digression to discuss contracts that, although largely hypothetical, are potentially more efficient than those considered above.

2.4.1 First-best

It is clear from equation (7) that the first-best contract is one in which the seller commits to $\eta = 1$. This contract is tantamount to one in which the price of groundwater is indexed (cf. Hart, 2009) to the seller’s post-transfer marginal product $f'(\frac{w-\tau}{a})$. Perhaps the complexity of this pricing scheme, or the lack of third-party state verification, renders such an arrangement impractical. At any rate, we do not observe anything like it in practice.\footnote{In Jacoby et al. (2004), borewell owners can commit to charging a groundwater price equal to the marginal extraction cost by having the water-buyer as a share-tenant. Sharecropping, however, is much less common in India than in Pakistan.}

Alternatively, a well-owner could achieve the first-best allocation by subsidizing the buyer’s planting cost at a rate of $1 - \eta$; obviously, allowing the planting investment to be contractible obviates the holdup problem. As a practical matter, however, it may be quite difficult for the seller to ensure the optimal ex-post demand for his water through such an incentive scheme if the buyer is free to adjust the intensity of cultivation. While this type of moral hazard problem, strictly speaking, lies outside of our model (because we have assumed that land and inputs like seed are always used in fixed proportions), it may explain the absence of such planting subsidies in our setting.

2.4.2 Mixed

Thus far, we have analyzed each type of contract in isolation, not allowing borewell owners to engage in both simultaneously. Our main reason for doing so is empirical; groundwater sales to multiple buyers under different contracts are rare in our data. The analysis of a mixed contract, however, is straightforward. Given his residual water $w - \tau C$ available ex-post, the borewell owner sells an amount $\tau P(\tau C)$ on a per-irrigation basis to buyer B. Working backwards, the amount sold to buyer A on a seasonal contract is the $\tau C$ that maximizes $a \left\{ Ef(\frac{w-\tau C-\tau P(\tau C)}{a}) - c \right\} + p\tau C$, subject to the participation and incentive constraints.
Clearly, the mixed contract does not achieve the first-best. While the allocation of water between the seller’s plot and that of buyer B is ex-post efficient, this is not the case for buyer A. Indeed, neither the ex-post nor the ex-ante distortion is entirely eliminated.\textsuperscript{16}

2.5 Characterizing the tradeoff

Returning to the main argument, we have already seen that the distortion induced by the long-term contract disappears when groundwater supply becomes perfectly certain, whereas the distortion induced by the spot contract does not. Next, we establish a general result about the dominance of long-term over spot contracts in our environment.

Recall that increases in $r$ correspond to mean preserving increases in uncertainty, with $r = 0$ indicating perfect certainty. Let $V_j(r)$ be the surplus derived from contract of type $j = C, P$,\textsuperscript{17} and note that $V_P(r, \eta)$ also depends on the bargaining weight $\eta$.

**Proposition 2 (Dominance)** If $g$ is strictly concave and $\tau_C(0) < w_L$,\textsuperscript{18} then (a) for some $\eta$, $\exists$ a unique $r^*(\eta)$ such that $V_C(r^*) = V_P(r^*, \eta)$; (b) $[V_C(r) - V_P(r, \eta)](r^* - r) > 0$.\textsuperscript{19}

Simply put, under the conditions of proposition 2, there can be a level of uncertainty at which the parties are indifferent between seasonal and per irrigation arrangements. If so, then the seasonal contract must dominate at low levels of uncertainty and per-irrigation sales at high levels of uncertainty.

Figure 1 illustrates the intuition underlying proposition 2, showing how the economic surplus generated by a borewell varies with uncertainty level $r$ under alternative water transfer arrangements. Regardless of arrangement, surplus always decreases with $r$ (see Appendix). In the case of autarky ($A$), in which the borewell irrigates exactly plot area $a$, surplus is

\textsuperscript{16}In the spirit of contracts as reference points, a buyer under a seasonal contract is precluded from selling back water to the borewell owner, or to anyone else, on a per-irrigation basis as this would presumably aggrieve the borewell owner (i.e., the buyer is not “entitled” to profit in this manner).

\textsuperscript{17}For the seasonal contract, surplus is given by the private returns to the well-owner; since the $PC$ is binding, the water-seller gets all the surplus. By contrast, in the per-irrigation case, we must consider the joint surplus of well-owner and water-buyer. It might be argued that the choice of per-irrigation over alternative arrangements should be governed by the water seller’s private returns as well. 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 private return would be resolved through side-payments.

\textsuperscript{18}In words, this latter condition states that the water transfer under perfect certainty is less than total water available in the worst state of the world. Otherwise, $V_C$ has a discontinuity at $r = 0$; i.e., at $r = \epsilon$, the optimal transfer must be discretely less than $\tau_C(0)$. In this case, $r^*(\eta)$ still exists for some $\eta$ but it is not necessarily unique. Part (b) of the proposition continues to hold, however, with respect to the largest $r^*$.

\textsuperscript{19}Proof: See Appendix.
\[ V_A = aE[f(w/a) - c] \]. \( V_A \) must lie strictly below first-best surplus \( V_U \) except at \( r = r_U \); at this level of uncertainty, \( \ell_U = a \) and autarky is the optimal unconstrained choice. When the borewell owner sells water under a seasonal contract, surplus \( V_C \) is also less than first-best (except under perfect certainty), coinciding with \( V_A \) at some positive level of uncertainty \( r_C < r_U \). Note that \( V_C \) declines relatively rapidly with \( r \) because higher uncertainty operates upon two margins under a seasonal contract: It leads to greater ex-post misallocation of groundwater across plots as well as to a contraction of overall area irrigated by the borewell (precautionary planting). Only the latter effect is operative under the per-irrigation arrangement. In this case, surplus \( V_P \) approaches \( V_U \) as \( \eta \) approaches one. Moreover, at some low level of bargaining power \( \eta = \bar{\eta} \), \( \ell_P = 0 \) and \( V_P \) coincides with \( V_A \). So, for some range of \( \eta \in (\bar{\eta}, 1) \), \( V_P \) and \( V_C \) must cross. Given such a crossing (at \( r^* \)), \( V_P \) coincides with \( V_A \) at a level of uncertainty \( r_P \) between \( r_C \) and \( r_U \). This shows that the spot contract can only dominate the long-term contract when uncertainty is high.

3 Data and Background

Our data come from a randomly selected survey of 2,423 borewell owners undertaken in six districts of Andhra Pradesh (AP) and Telangana (until 2014, also part of 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 774 borewell owners were followed up from that study’s 2010 household survey. Guntur and Kadapa 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.\textsuperscript{20}

The 1,649 borewells in Guntur, Kadapa, East and West Godavari form our estimation sample, whereas the 774 borewells in more arid Anantapur and Mahabubnagar districts constitute a holdout sample that we use for model validation and counterfactual analysis.

\textsuperscript{20}A total of 144 villages were covered (21-25 per district) in the survey. Our sample is broadly representative of areas with sufficient groundwater for \textit{rabi} cultivation and where groundwater is the sole source of irrigation during that season (villages in canal command areas were avoided).
3.1 Recharge and uncertainty

As in much of India, farmers in AP rely almost exclusively on groundwater during the rabi (winter or dry) 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.\(^{21}\) 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 is explained by the limited storage capacity of the shallow hard rock aquifers that characterize the region. Most 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).\(^{22}\)

This annual cycle of aquifer replenishment and draw-down throughout AP is central to our analysis of groundwater markets. Although 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 measure the degree of uncertainty, as part of the borewell owner’s survey we fielded a well-flow expectations module, 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).\(^{23}\) 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 2 shows the distributions of groundwater uncertainty (well-specific coefficients of variation of end-of-season flow) in both the estimation and holdout samples. The first thing to notice is that virtually no borewell owner (except five in the estimation sample) reports having a perfectly certain supply of groundwater. Secondly, the difference across samples is striking; uncertainty is much higher in the holdout districts of Anantapur and Mahabubnagar where aquifer recharge is relatively meager. One goal of the empirical work

\(^{21}\)Appendix figure A.1 documents the rising importance of borewell irrigation in all of India, in AP as a whole (prior to the 2014 partition), and in the six districts of our survey. See also World Bank (2005).

\(^{22}\)Shah (2009) emphasizes the role of different aquifer types in shaping groundwater governance arrangements in South Asia.

\(^{23}\)To appreciate how discharge can be fractional for an extended period, the metaphor for the aquifer to keep in mind is that of a sponge rather than of a bathtub.
below is to investigate whether this difference in uncertainty levels can account for the similarly dramatic divergence in groundwater market activity across regions.

### 3.2 Land fragmentation, fixed costs, and groundwater markets

A second crucial element of our analysis 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 land market.\(^{24}\) In our data, nearly 80 percent of plots were acquired through inheritance.

Land fragmentation would be irrelevant, of course, were groundwater markets frictionless. If so, 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.\(^{25}\) Neither implication of frictionless groundwater markets, however, is consistent with our data.

Our survey covers around 9600 plots, each of which either has a borewell itself or is adjacent to a plot that does and, thus, could in principle receive a transfer of groundwater. Figure 3 shows that borewells are actually much less likely on small plots than on large ones. One might think that a random allocation of borewells across space could generate such a pattern mechanically; larger plots would be more likely to have borewells insofar as they constitute the majority of farmland area. But this ignores the fact that well placement is determined by individual decision-makers at the plot-level. If there is an equal probability of successfully finding groundwater regardless of where one drills for it and each plot-owner makes the same number of drilling attempts, then the likelihood of observing a borewell should be equal across plots, regardless of size.

To be sure, owners of small plots may also be less wealthy and thus unable to afford multiple drilling attempts, or any attempts at all for that matter (see, e.g., Fafchamps and Pender, 1996). To control for wealth, 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. We then partial out the effect of wealth (as proxied by total landholdings) using dummies for each of the deciles of total landownership. The resulting figure 4 continues to show an

\(^{24}\) 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).

\(^{25}\) This is just an application of the Separating Hyperplane Theorem, by which asset ownership is irrelevant for production decisions, or, if one prefers, the strong Coase Theorem.
increasing borewell probability as plot size increases. Finally, figure 5 shows that small plots are much more likely to be left fallow in the dry season than large plots. Taken together, this evidence indicates frictions in groundwater markets.

3.3 Adjacency approach

To capture transfers of groundwater, which typically occur between adjacent plots so as to minimize conveyance losses,\textsuperscript{26} we departed from the usual household-based sampling strategy. Instead, each of the 2307 randomly chosen respondents (borewell owners) was also asked to report on all the plots adjacent to the one containing the reference borewell,\textsuperscript{27} including characteristics of the landowner, details on how the plot was irrigated during the \textit{rabi}, if not left fallow, and on the transfer arrangement if one occurred. This \textit{adjacency} approach provides information about not only the transfers that did happen but also about those that could have happened but did not.

The number of plots adjacent to a reference borewell varies from 1 to 7, with a mode of 3. Table 2 provides descriptive statistics on these plots, showing that in the estimation sample 42\% are irrigated in whole or in part by the reference borewell. This figure falls to just 15\% of adjacent plots in the holdout sample, even though the proportion owned by brothers, who are more likely to be co-owners of the reference borewell,\textsuperscript{28} is substantially higher in the holdout sample. A similarly wide disparity exists in the percentages of plots accessing other borewells, which may or may not be in the adjacency. Thus, transfers of groundwater, and especially sales,\textsuperscript{29} are comparatively limited in the two holdout districts and commensurate with this is a much higher fraction of fallow plots in \textit{rabi} season.

In figure 6, we pool the two samples and aggregate plot-specific data to the adjacency level. The top curve shows how borewell density in the adjacency—the average of the plot-level borewell indicator used in figure 3 weighted by the plot’s area share in the adjacency—varies with average plot size. Mirroring figure 3 at the adjacency level, it shows that borewell

\textsuperscript{26}Most irrigation water is transferred through unlined field channels with high seepage rates. While 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.

\textsuperscript{27}There are 114 cases of the reference plot having a second borewell and 2 cases of it having a third borewell, which gives a total of 2423 reference wells.

\textsuperscript{28}In constructing the data set for the structural estimation, we merge the plots of all co-owners of the reference borewell found in the adjacency.

\textsuperscript{29}As most sales transactions (86\%) are between non-relatives, holdup concerns are not \textit{prima facie} misplaced. Among these non-relative transactions, 60\% are between members of the same caste. However, 57\% of non-related adjacent plot owners are of the same caste as the owner of the reference borewell. Hence, Anderson’s (2011) suggestion of caste-based barriers to groundwater trade does not find support in our data.
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 its co-owners). The frequency of such transactions falls with average plot size in the adjacency. So borewell density and groundwater market activity are substitutes, both driven by the degree of local land fragmentation.

A final point concerns the nature of competition in groundwater markets. For convenience, our theoretical analysis assumes bilateral monopoly. Our data, however, indicate that the median number of borewells in an adjacency is 2. Given their dispersion across space and the high cost of moving water, we may think of water buyers and sellers as being connected in trading networks. Corominas-Bosch (2004) provides a model of bilateral bargaining in such networks, which could be applied to spot transactions in our setting. Her key result is that a buyer’s bargaining power (η in our case) depends not only on the relative number of buyers and sellers, but also on the link structure of the network. For this reason, we incorporate unobserved heterogeneity in η into our structural econometric model.

### 3.4 Precautionary planting and risk aversion

Proposition 1 shows that uncertainty in groundwater availability can lead to precautionary planting. This, however, hinges on the properties of \( g \), the marginal return to planting, which is not directly observable. To motivate our functional form assumptions, we now present a reduced-form analysis of planting decisions.

Crops grown during *rabi* season in AP fall 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. Since a field that, planted to ID crops, would take 3 days to irrigate, would take a week to irrigate under wet crops, we use the equivalence 1 acre wet = \( \frac{7}{3} \) acre ID to compute total area irrigated by a borewell.\(^{31}\)

---

\(^{30}\)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).

\(^{31}\)We carry over this efficiency units assumption to our structural estimation as well for reasons of tractability. Consider, briefly, the practical obstacles to explicitly incorporating the choice between separate wet and ID crop technologies. In the first place, conditional on area choices of each crop type, farmers would presumably allocate groundwater ex-post across crops in response to the realized \( w \). This gives rise to one additional optimality condition for each state of nature. Second, there would be two cultivated area choices, and farmers are observed opting for mixed wet-ID cropping as well as for monoculture of either type. To rationalize the data, our structural model would need two error terms (instead of just one, as currently assumed) and would have to account for the two types of corner solutions in cropped area. Third, for any form of groundwater transfer, each cell of the \( 3 \times 3 \) matrix of wet-ID-mixed cropping decisions of borewell
We investigate the effect of uncertainty on *rabi* area irrigated by a borewell, conditional on two alternative area measures. First, a *hypothetical* measure of area 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.*”\(^{32}\) If there is a precautionary planting motive, we expect borewell owners would irrigate more area under certainty than they actually irrigate. This is, in fact, what we see, with a mean difference between actual and hypothetical areas of about 30% (median 24%). To see how such (hypothetical) precautionary behavior varies with groundwater uncertainty, we regress log actual area irrigated by a borewell on log hypothetical area irrigated under certainty, the log($CV_i$) of end-of-season flow (see figure 2), as well as the log of outlet pipe-width (to control for borewell capacity). The negative coefficient on log($CV_i$) in column 1, panel (a) of Table 3, indicates that, as groundwater uncertainty increases, borewell owners deviate further (by irrigating less area) from their perfect certainty benchmark. This result is robust to controls for additional borewell characteristics (pump horse-power, log well depth, number of nearby wells, presence of groundwater recharge) in column 2.

Next, we consider planting behavior relative to the *autarky* benchmark, which is to say irrigating exactly the borewell-plot area. Thus, we run the same regressions as before, but this time condition on log of reference plot area. A total area irrigated less than reference plot area (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. Results for these regressions, reported in panel (b) of Table 3, show that again higher groundwater uncertainty leads to a reduction in area planted.

While this evidence supports a precautionary planting motive, we have not yet established the theoretical mechanism.\(^{33}\) 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

\(^{32}\)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*.

\(^{33}\)We should also note that it is only by estimating a structural model that we can distinguish precautionary planting per se from the effects of contracting distortions and other transactions costs.
towards risk collected from well owners in the survey. The first measure, \( RISK_1 \), 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), \( RISK_2 \) 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 \( RISK_2 \) indicate greater risk aversion. In columns 3 and 4 of Table 3, we report results of adding, respectively, \( RISK_1 \) and \( RISK_2 \), and, most importantly, their interactions with \( \log(CV_i) \), in the corresponding baseline regression of column 2. The estimated coefficients on these interactions, and particularly their lack of significance, betrays no indication that highly risk averse borewell owners are especially responsive to groundwater uncertainty. Precautionary planting, therefore, does not appear driven by risk preference. For this reason, as noted in (A.4), we assume risk neutrality throughout the paper.

4 Estimation framework

Table 4 provides descriptive statistics for the estimation and holdout samples according to the groundwater transfer choices made by the borewell owner.\(^{34}\) Two-thirds of the estimation sample transferred groundwater to other plots in the adjacency during the past rabi season, most often through a seasonal sales contract, followed by per-irrigation sales and, very far behind, by leasing. 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 is substantially lower for the per-irrigation arrangement.

By contrast, transactions are virtually nonexistent in the holdout sample, consisting of borewells in drought-prone Anantapur and Mahabubnagar districts, where groundwater is both relatively scarce and uncertain. Particularly stark is the collapse of the seasonal contracts as compared to its prevalence in the estimation sample. Also, among those borewell owners not transferring groundwater, a much higher fraction are unconstrained in the holdout sample (62%) than in the estimation sample (34%); in other words, fewer would be willing to make a transfer even in the absence of transactions costs.

\(^{34}\)Four borewells from the original 2,423 were dropped due to missing characteristics. Also, 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.
4.1 Leasing

To account for leasing, we assume that this arrangement entails an efficiency cost making it less attractive than irrigating one’s own land. A 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 land leasing would be problematic (in the next subsection, we also introduce a fixed cost of leasing).

Thus, let \( \gamma > 0 \) be the proportional increment to cultivation costs that applies only to leased land. Optimal leased area is then given by

\[
\ell_L = \arg \max (a + l) \{ E_f(w/(a + l)) - c \} - \gamma cl. \tag{8}
\]

4.2 Fixed transactions costs

To explain why well-owners remain in autarky (choice A) rather than transfer tiny amounts of water 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. On the one hand, a lease or water sale will be difficult if all neighboring plots are already irrigated by their own borewell. On the other hand, there may be incentives for farmers to drill wells of their own in areas less conducive to such markets. To capture heterogeneity in \( \kappa \) across adjacencies, therefore, we also need, in effect, an instrument for borewell density.

As figure 6 suggests, we do have such an instrument: average plot size in borewell \( i \)'s adjacency, \( \bar{a}_i \), which is plausibly unrelated to groundwater markets except through its strong positive association with borewell density. Thus, we incorporate heterogeneity in irrigable land availability by introducing \( \bar{a}_i \) 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 is willing to forgo dry-season cultivation.\(^{35}\)

Putting these two elements together, we have

\[
\kappa_{ij} = \pi_j \exp(\beta \bar{a}_i) \tag{9}
\]

\(^{35}\)It may also be more difficult to lease a fraction of a neighboring plot than it is to irrigate an equivalent area through selling water to that plot’s owner-irrigator.
for water transfer type \( j \), where \( \bar{K}_L \neq \bar{K}_C = \bar{K}_P \equiv \bar{K}_T \). In particular, we expect that \( \bar{K}_L > \bar{K}_T \).

### 4.3 Functional form

We next assume that the intensive production function, \( f \), takes the form

\[
    f(\omega) = \zeta \omega^{\alpha}.
\]

where \( 0 < \alpha < 1 \). That is, \( F \) is Cobb-Douglas in land and water. Henceforth, we normalize the parameter \( \zeta = 1 \), which, given the homogeneity of \( f \), is equivalent to fixing the units of output and, thus, of cultivation costs \( c \). With these assumptions, the implied \( g \) (marginal return to planting) is globally concave. Therefore, by proposition 1 and consistent with the empirical evidence presented in Section 3, there is a precautionary planting motive.

Combining (10) with (1) yields

\[
    \ell_U = \left( \frac{1 - \alpha}{c} Ew^\alpha \right)^{1/\alpha}
\]

and with definition 2 delivers

\[
    V_U = \frac{\alpha c}{1 - \alpha} \ell_U.
\]

Solutions for area irrigated and economic surplus under the remaining arrangements are reported in Table 5. A closed form for area irrigated is lacking only in the case of the seasonal contract. In the case of per-irrigation sales, we find that the ratio of economic surplus to unconstrained surplus, \( V_P/V_U \), is simply equal to a constant \( \delta < 1 \) (defined in Table 5). Thus, \( 1 - \delta \) represents the relative distortion of the per-irrigation arrangement, which vanishes as the buyer’s bargaining weight \( \eta \) approaches unity.

### 4.4 Borewell discharge

As noted in Section 3, well-owners report conditional probabilities for five water flow states, corresponding to “full”, 3/4, 1/2, 1/4, and zero flow. For empirical purposes, therefore, the groundwater distribution \( \psi(w) \) is discrete, consisting of five points of support \( k = 0, 1, 2, 3, 4 \) and corresponding borewell-specific probabilities, \( \pi_{ki} \).

Since water discharge is proportional to the square of pipe radius \( R_i \), we have

\[
    w_{ki} = \lambda R_i^2 k
\]
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 $\frac{1}{4}$.

So, expected borewell discharge $Ew_i = R^2 \sum_k \pi_{ki} k/4$ varies with pipe-width as well as with expected flow. For example, in the estimation sample the average pipe-width is 3.0 (inches) and the average expected flow is 0.90, as compared to 2.1 and 0.70, respectively, in the holdout sample. Thus, the typical borewell in the estimation sample differs from that in the holdout sample both in having a lower second moment of discharge, as noted earlier, and in having a higher first moment (by a factor of nearly 3).

### 4.5 Cost disturbance

To explain why different water transfer arrangements (including none at all) are chosen across observationally equivalent borewells, as well as why different areas are cultivated conditional on the transfer arrangement, we introduce a cost disturbance $\varepsilon_i \sim N(0, \sigma^2_\varepsilon)$ such that $c_i = \tau_e^{\varepsilon_i}$. We assume cost heterogeneity reflects variation in local (adjacency-level) conditions, such as soil texture, depth, and water-retention capacity, or in the shadow price of inputs like seed and fertilizer, rather than in cultivator characteristics.

Inserting $c_i$ into the expressions given in the first column of Table 5 and inverting the resulting functions $\ell_{ij} = \ell_j(\varepsilon_{ij})$ yields choice-specific residuals $\varepsilon_{ij}$ for $j = U, L, C, P$, where $\ell_{ij}$ is the area under that arrangement for borewell $i$.

### 4.6 Measurement error

To allow flexibility in fitting variances of these irrigated areas, we introduce a normally distributed measurement error $u_i \sim N(0, \sigma^2_u)$ such that

$$\ell_{ij}^* = \ell_{ij} e^{u_i}.$$  \hspace{1cm} (14)

where the asterisk denotes observed quantities. For the simplest case, $j = U$, equations (11) and (14) deliver

$$\log \ell_{iU}^* = \frac{1}{\alpha} \left[ \log(1 - \alpha) - \log \bar{c} + \log \sum_k \pi_{ki} w_{ki} - \varepsilon_{iU} \right] + u_i.$$ \hspace{1cm} (15)

Thus, $\text{var}(\log \ell_{iS}^*)$ is not pinned down by the structural error variance alone.

For choices $j = L, C, P$, the error terms $\varepsilon$ and $u$ no longer appear additively as they do in equation (15). Since the branches of the likelihood function corresponding to these choices
involve Monte Carlo integration, we relegate the discussion of estimation with measurement error to an Appendix.

4.7 Unobserved parameter heterogeneity

The buyer's bargaining weight $\eta$ is a key parameter determining the choice of seasonal contract ($C$) versus per-irrigation ($P$) arrangement. But, as discussed in Section 3, $\eta$ depends on the link structure of the buyer-seller network within each adjacency and possibly across overlapping adjacencies. Since we did not collect data on these link structures (nor would it have been an easy task), we treat the true $\eta$ 'type' as unobserved.

Assume, therefore, $M$ types, where

$$\eta = \sum_{m=1}^{M} I(\text{type} = m) \eta_m$$ (16)

We estimate each $\eta_m$ along with the corresponding type probabilities $\rho_m$.

4.8 Likelihood function

The mixed continuous/discrete choice likelihood function involves the density of $\varepsilon_{ij}$ and the probabilities of choices $j = U, A, L, C, P$. The latter are determined by thresholds

$$V_A(\varepsilon_{jA}^*) = V_j(\varepsilon_{jA}^*) - I(j \neq U)\kappa_{ij} \quad j = U, L, C, P$$

$$V_j(\varepsilon_{jk}^*) - \kappa_{ij} = V_k(\varepsilon_{jk}^*) - \kappa_{ik} \quad (j, k) = \{(L, C), (L, P), (P, C)\}$$, (17)

defined by the crossings of the value functions on the support of the cost disturbance. The first set of equalities in (17) yield the cultivation costs at which the well-owner is just indifferent between autarky and transferring water under each of the three respective arrangements, whereas the second set of equalities give the cost thresholds between alternative transfer arrangements. Figure 7 illustrates one of many partitions of the support of $e^\varepsilon$ into regions where either per-irrigation, leasing, seasonal contracts, autarky, or unconstrained self-cultivation uniquely provide the highest surplus.

Let $Pr(j|\Theta_m, Z_i)$ denote the probability of choice $j$ conditional on the parameters of the model $\Theta_m = (\alpha, \eta_m, \gamma, c, \overline{c}, \kappa_L, \kappa_T, \beta)$ for type $m$ and the data $Z_i = (\pi_{ki}, R_i^2, a_i, \overline{a}_i)$. Given (17),
we have

\[
\Pr(U|\Theta_m, Z_i) = 1 - \Phi(\varepsilon_{imUA}/\sigma_\varepsilon)
\]

\[
\Pr(A|\Theta_m, Z_i) = \Phi(\varepsilon_{imUA}/\sigma_\varepsilon) - \Phi(\tilde{\varepsilon}_{imA}/\sigma_\varepsilon)
\]

\[
\Pr(j|\Theta_m, Z_i) = \sum_r \left[ \Phi(\varepsilon_{imjr}/\sigma_\varepsilon) - \Phi(\tilde{\varepsilon}_{imjr}/\sigma_\varepsilon) + \Phi(\varepsilon_{imjr}^{(2)}/\sigma_\varepsilon) - \Phi(\tilde{\varepsilon}_{imjr}^{(2)}/\sigma_\varepsilon) \right]^{D_{ir}} \quad j = L, C, P
\]

where \( \Phi \) is the normal cdf, \( \tilde{\varepsilon}_{imA} = \max \{ \varepsilon_{imLA}^{*}, \varepsilon_{imCA}^{*}, \varepsilon_{imPA}^{*} \} \), the \( \varepsilon_{imjr}^{(n)} \) and \( \tilde{\varepsilon}_{imjr}^{(n)} \) are, respectively, upper and lower limits of the regions of integration for the probability of choice \( j \) for type \( m \) 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 type-specific value function crossings \( \varepsilon_{jmA}^{*} \) and \( \varepsilon_{jmk}^{*} \).\(^{36}\) Note that for \((j, k) = \{(L, C), (P, C), (L, P)\}\) there can be two value function crossings, enumerated here by the index \( n \). Also, for well-owners in autarky \((j = A)\), we do not know the transfer arrangement they would have made. Hence, the lower limit of integration for their choice probability, \( \tilde{\varepsilon}_{imA} \), represents the highest \( \varepsilon_i \) that would have induced them to make any kind of water transfer.

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

\[
\mathcal{L}_i(\Theta_m) = \Pr(A|\Theta_m, Z_i)^{d_{iA}} \prod_{j \neq A} \left[ \Pr(j|\Theta_m, Z_i) \frac{\alpha}{\sigma_\varepsilon} \phi\left( \frac{\varepsilon_{ij}}{\sigma_\varepsilon} \right) \right]^{d_{ij}},
\]

where \( \phi \) is the standard normal density. As noted, we incorporate measurement error in the Appendix, but this only affects the continuous part of the likelihood, not the choice probabilities. Finally, taking the probability-weighted average across types yields unconditional log-likelihood

\[
\log \mathcal{L}(\Theta) = \sum_i \log \left[ \sum_m \rho_m \mathcal{L}_i(\Theta_m) \right].
\]

\(^{36}\)In some regimes, \( j \) may not be optimal for any \( \varepsilon_i \), in which cases we penalize the likelihood by setting the choice probability to a very small number.
4.9 Identification

The nonlinear parameter restrictions embedded in the thresholds defined by (17) identify the structural parameters and $\sigma_\varepsilon$ through the discrete choices alone. However, the continuous choices of areas irrigated, the $\ell_{ij}$, are an additional source of identification. To see this, return to the example of equation (15) for the case $j = U$ and note that $\alpha$ controls how fast irrigated area falls with variability in water supply and $\bar{\sigma}$ is pinned down by average irrigated area. Similarly, information on $\ell_{iP}$ and $\ell_{iL}$ would over-identify $\eta_m$ and $\gamma$, respectively. So, there is a potential efficiency gain from estimating the joint continuous/discrete model.

5 Results

5.1 Parameter estimates

Parameter estimates for the model are reported in Table 6 along with standard errors. The estimates appear reasonable and extremely precise. In particular, $\alpha$ is considerable less than one, whereas a value close to one would have implied little role for groundwater uncertainty. The marginal inefficiency of leasing land, $\gamma$, is virtually zero, which means that the paucity of lease activity in the data is largely driven by the high fixed transactions cost, $\kappa_L$ (relative to $\kappa_T$). As expected, the coefficient $\beta$ on average plot size in the adjacency is positive, indicating that owners of borewells surrounded by larger plots (i.e., those more likely to have borewells of their own) face higher transactions costs of arranging groundwater transactions.

Our estimates $\hat{\eta} = 0.935$ and $\hat{\alpha} = 0.218$ translate (cf., definition of $\delta$ in Table 5) into a 3.1% efficiency loss due to holdup in the per-irrigation arrangement. The modesty of this distortion suggests that water buyer and seller might be engaged in repeated interactions, perhaps over multiple seasons, rather than in the one-shot bargaining game that we assume in our stylized model.

5.2 Model fit

To assess model fit in the estimation sample, we predict for each borewell owner a probability for each choice $j = U, A, L, C, P$. Also, for those borewell owners whose choice is either $j = C$ or $P$, we compute the expected transfer in terms of irrigated area relative to own area as

$$E[\log(1 + \ell_{ij}/a_i)] = \int \log(a_i + \ell_j(\varepsilon_{ij}))d\tilde{\Phi}_{ij} - \log a_i,$$  \hspace{1cm} (21)
where $\tilde{\Phi}_{ij}$ is the appropriate truncated normal cdf of $\varepsilon$. For $j = P$ this expression is linear in $\varepsilon_{iP}$. However, for $j = C$ we must use Monte-Carlo integration by drawing $\varepsilon_{iC}$ from $d\tilde{\Phi}_{iC}$, calculating the corresponding $\ell_C(\varepsilon_{iC})$, and averaging the integrand across draws.

Finally, we construct for every borewell owner, regardless of choice, a measure of predicted groundwater market activity relative to own area irrigated as

$$E[\log(1 + \ell_{iT}/a_i) = Pr_i(C)E[\log(1 + \ell_{iC}/a_i)|C] + Pr_i(P)E[\log(1 + \ell_{iP}/a_i)|P]$$

(22)

where the $Pr_i(j)$ are borewell-specific predicted choice probabilities.

Column 2 of Table 7 reports the results of these calculation in the estimation sample.\textsuperscript{37} The model does a decent job of predicting borewell owners’ choices of transfer arrangements ($L$, $C$, and $P$), but much less well on the $U$-$A$ margin. Evidently, the reason for this is that our likelihood function maximization is not just looking to match choice probabilities, but also to fit area irrigated under each arrangement. The last row of the table shows that the model indeed does reasonably well relative to the data in predicting $E[\ell_{iT}]$.

Also in Table 7, we report predictions of our model (using parameters estimated on the estimation sample) for borewells in the holdout sample. As one would hope, the model indicates much less groundwater market activity in the drought-prone districts, capturing well the virtual disappearance of the seasonal contract. Remarkably, the model fits the probability of being unconstrained ($P(U)$) much better in the holdout sample than in the estimation sample, but it also predicts higher prevalence of per irrigation contracts than is otherwise found in the data.

### 5.3 Counterfactuals

Our first counterfactual asks what would happen to groundwater market activity in the holdout districts if uncertainty is reduced to that prevailing in the estimation districts. To perform this exercise, for each borewell in the holdout sample we draw a 5-tuple of $\pi_{ki}$ at random (with replacement) from the estimation sample and use it instead of the actual vector of probabilities, rescaling $\lambda$ (cf. equation 13) so as to keep $\bar{w}_i$ constant. The analog is also run by assigning each borewell in the estimation sample a vector of $\pi_{ki}$ drawn at random from the holdout sample.

Results for this mean-preserving change in $\sigma_w$ using the estimation sample as the baseline

\textsuperscript{37}In this version of the paper, we do not perform the Monte Carlo integration over the conditional error distribution, but simply set the error term to zero for each borewell owner. Results in Table 7 are preliminary.
(column 3 of Table 7) and that using the holdout sample as the baseline (column 7) both show a dramatic shift between seasonal contract and per-irrigation arrangement, but little change in groundwater market activity ($E[\ell_{IT}]$). Thus, contractual choice appears to be far more responsive to uncertainty levels than is the overall willingness to trade water.

We also performed the analogous exercise of replacing $\bar{a}$ with a randomly drawn counterpart from the alternative sample. In contrast to the case of groundwater supply uncertainty, differences in transactions costs across regions, which in our model are driven by differences in $\bar{a}$, have negligible impacts on groundwater transfer arrangements (see cols. 4 and 8 of Table 7). In part, this reflects the small mean difference in average adjacency plot size across estimation and holdout samples (0.45 acres).

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 $V_U$ rather than $V_j - \kappa_j$ under the best alternative water arrangement.

On average, this hypothetical consolidation would lead to a 13.5% increase in irrigated area and a 7.2% increase in economic surplus from dry-season cultivation in the four estimation sample districts. Gains in the holdout sample districts would be negligible due to the lack of extant groundwater markets. In figure 8, we plot counterfactual changes in both irrigated area and surplus as a function of plot area adjusted for pipe-width (i.e., $a_i/r_i^2$). The gains to consolidation are decreasing in adjusted plot size, as borewell owners with larger plots are, of course, less compelled to engage in costly groundwater transactions.

6 Conclusion

We have developed a model of contractual arrangements, which, in the spirit of the transactions cost literature, features a tradeoff between ex-post and ex-ante inefficiency. Long-term contracts protect relationship-specific investment but are less flexible than spot contracts. We have shown formally that, as payoff uncertainty rises, there is a shift from long-term to spot contracting. Our structural estimates of the model using micro-data on area irrigated under different groundwater transfer arrangements reveal a quantitatively important role for uncertainty in shaping contractual form; less so in the development of groundwater markets. Given the distortions induced by this uncertainty, we find substantial gains from land consolidation in terms of increased irrigated area and surplus, especially for those borewell
owners with the smallest plots.

While the specific context of groundwater is, of course, unique, we believe that this paper has broader implications for how we think about markets in uncertain environments.
References


Figure 1: Long-term versus spot contracts and uncertainty
Figure 2: Borewell-level groundwater supply uncertainty

Note: Estimation sample drawn from East Godavari, Guntur, Kadapa, and West Godavari districts; holdout sample from drought-prone Anantapur and Mahabubnagar districts.
Figure 3: **Presence of a borewell and area of plot**

![Graph showing the proportion of plots with borewells against plot area (acres). The graph includes a local cubic polynomial fit and cumulative distribution function.](image)

Note: Conditional local cubic polynomial fit partials out dummies for the deciles of total land area owned.

Figure 4: **Presence of a borewell and area of plot controlling for wealth**

![Graph showing the proportion of plots with borewells against plot area (acres), with an additional line for plot area/total area owned.](image)

Note: Conditional local cubic polynomial fit partials out dummies for the deciles of total land area owned.
Figure 5: Dry-season fallow and area of plot
Figure 6: **Borewell density, groundwater transfers, and average plot size**
Figure 7: A POSSIBLE PARTITION OF THE ERROR-SPACE
Figure 8: Impact of hypothetical land consolidation
Table 1: Decisions and Timing

<table>
<thead>
<tr>
<th>Arrangement</th>
<th>Contract type</th>
<th>Ex-ante</th>
<th>Ex-post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal</td>
<td>long-term</td>
<td>$p, \tau, \ell$</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>surplus divided</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per-irrigation</td>
<td>spot</td>
<td>$\ell$</td>
<td>$p, \tau$</td>
</tr>
<tr>
<td></td>
<td>surplus divided</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: $p$ is the price per unit of irrigation, $\tau$ is the total transfer of groundwater, and $\ell$ is area irrigated by the buyer.

Table 2: Characteristics of Adjacent Plots

<table>
<thead>
<tr>
<th></th>
<th>Estimation Sample</th>
<th>Holdout Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean number of adjacent plots per adjacency</td>
<td>3.39</td>
<td>3.59</td>
</tr>
<tr>
<td>Mean area of adjacent plots (acres)</td>
<td>2.95</td>
<td>3.34</td>
</tr>
<tr>
<td>% of adjacent plots left fallow in rabi</td>
<td>2</td>
<td>38</td>
</tr>
<tr>
<td>% of adjacent plots irrigated in rabi by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• reference borewell</td>
<td>42</td>
<td>15</td>
</tr>
<tr>
<td>of which, % irrigated under</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- joint ownership of reference borewell</td>
<td>22</td>
<td>95</td>
</tr>
<tr>
<td>- land lease</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>- water sale</td>
<td>71</td>
<td>2</td>
</tr>
<tr>
<td>• own borewell</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>• other borewell</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>% of adjacent plots owned by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• brother of reference borewell owner</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>• other relative</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>• unrelated member of same caste</td>
<td>50</td>
<td>33</td>
</tr>
<tr>
<td>• unrelated person of different caste</td>
<td>32</td>
<td>35</td>
</tr>
<tr>
<td>Number of plots</td>
<td>4943</td>
<td>2350</td>
</tr>
</tbody>
</table>
Table 3: Precautionary Planting and Risk Aversion

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Hypothetical benchmark (N = 2,410)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(area irrigated under certainty)</td>
<td>0.950***</td>
<td>0.939***</td>
<td>0.939***</td>
<td>0.939***</td>
</tr>
<tr>
<td></td>
<td>(0.0204)</td>
<td>(0.0213)</td>
<td>(0.0213)</td>
<td>(0.0213)</td>
</tr>
<tr>
<td>log(pipe width)</td>
<td>-0.0710**</td>
<td>-0.0682**</td>
<td>-0.0667**</td>
<td>-0.0685**</td>
</tr>
<tr>
<td></td>
<td>(0.0290)</td>
<td>(0.0270)</td>
<td>(0.0272)</td>
<td>(0.0271)</td>
</tr>
<tr>
<td>log(CV)</td>
<td>-0.125***</td>
<td>-0.105***</td>
<td>-0.161***</td>
<td>-0.0918*</td>
</tr>
<tr>
<td></td>
<td>(0.0249)</td>
<td>(0.0248)</td>
<td>(0.0611)</td>
<td>(0.0515)</td>
</tr>
<tr>
<td>$RISK^1$</td>
<td>-0.0155</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0237)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(CV) $\times RISK^1$</td>
<td>0.00904</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00929)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$RISK^2$</td>
<td>-0.201</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(CV) $\times RISK^2$</td>
<td>-0.0888</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0953)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.759</td>
<td>0.763</td>
<td>0.763</td>
<td>0.763</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Autarky benchmark (N = 2,411)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(borewell plot area irrigated)</td>
<td>0.568***</td>
<td>0.555***</td>
<td>0.556***</td>
<td>0.556***</td>
</tr>
<tr>
<td></td>
<td>(0.0149)</td>
<td>(0.0145)</td>
<td>(0.0144)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>log(pipe width)</td>
<td>0.467***</td>
<td>0.453***</td>
<td>0.451***</td>
<td>0.452***</td>
</tr>
<tr>
<td></td>
<td>(0.0365)</td>
<td>(0.0405)</td>
<td>(0.0406)</td>
<td>(0.0406)</td>
</tr>
<tr>
<td>log(CV)</td>
<td>-0.568***</td>
<td>-0.536***</td>
<td>-0.509***</td>
<td>-0.478***</td>
</tr>
<tr>
<td></td>
<td>(0.0282)</td>
<td>(0.0289)</td>
<td>(0.0757)</td>
<td>(0.0618)</td>
</tr>
<tr>
<td>$RISK^1$</td>
<td>-0.0155</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0237)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(CV) $\times RISK^1$</td>
<td>0.00431</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$RISK^2$</td>
<td>-0.201</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(CV) $\times RISK^2$</td>
<td>-0.0888</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0953)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.625</td>
<td>0.634</td>
<td>0.635</td>
<td>0.635</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors in parentheses adjusted for clustering on borewell (*** p<0.01, ** p<0.05, * p<0.1). Dependent variable in all regressions is log(total irrigated area by borewell); 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 4: Descriptive Statistics

<table>
<thead>
<tr>
<th>Choice</th>
<th>Estimation Sample</th>
<th>Holdout Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$\ell_j$</td>
</tr>
<tr>
<td>$U =$ unconstrained</td>
<td>184</td>
<td>6.64</td>
</tr>
<tr>
<td></td>
<td>(6.00)</td>
<td>(10.33)</td>
</tr>
<tr>
<td>$A =$ autarky</td>
<td>362</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>(15.16)</td>
</tr>
<tr>
<td>$L =$ leasing</td>
<td>97</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td>(3.78)</td>
<td>(3.24)</td>
</tr>
<tr>
<td>$C =$ seasonal contract</td>
<td>642</td>
<td>3.11</td>
</tr>
<tr>
<td></td>
<td>(2.73)</td>
<td>(7.05)</td>
</tr>
<tr>
<td>$P =$ per-irrigation sale</td>
<td>361</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>(1.63)</td>
<td>(2.72)</td>
</tr>
<tr>
<td>Total</td>
<td>1646</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(9.45)</td>
<td>(5.74)</td>
</tr>
</tbody>
</table>

**Notes:** Sample means (standard deviations).

### Table 5: Theoretical Solutions for Irrigated Area and Economic Surplus

<table>
<thead>
<tr>
<th>Choice $j$</th>
<th>Area ($\ell_j$)</th>
<th>Surplus ($V_j$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U =$ unconstrained</td>
<td>((\frac{1-a}{c} E w^\alpha)^{1/\alpha})</td>
<td>(\frac{ac}{1-a} \ell_U)</td>
</tr>
<tr>
<td>$A =$ autarky</td>
<td>$a$</td>
<td>$a^{1-\alpha} E w^\alpha - ca$</td>
</tr>
<tr>
<td>$L =$ leasing</td>
<td>((1 + \gamma)^{-1/\alpha} \ell_U - a)</td>
<td>((1 + \gamma)^{1-1/\alpha} V_U + \gamma ca)</td>
</tr>
<tr>
<td>$C =$ seasonal contract</td>
<td>$\ell_C$ solves $E \Omega^{\alpha-1} = 1$</td>
<td>(\frac{ac}{1-a} [E \Omega^{\alpha} + \alpha \ell_C/a - (1 - \alpha)])</td>
</tr>
<tr>
<td>$P =$ per-irrigation sale</td>
<td>$\eta^{1/\alpha} \ell_U - a$</td>
<td>$\delta V_U$</td>
</tr>
</tbody>
</table>

**Notes:** \(\Omega = \frac{1}{\alpha} \left[ w \left( \frac{1-a}{c} \right)^{1/\alpha} - \ell_C \right]\) and \(\delta = \frac{1-\eta^{(1-\alpha)}}{\alpha} \eta^{1/\alpha-1}\).
### Table 6: Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.218</td>
<td>0.0307</td>
</tr>
<tr>
<td>$\bar{c}$</td>
<td>0.731</td>
<td>0.0371</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.001</td>
<td>0.0032</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.935</td>
<td>0.0096</td>
</tr>
<tr>
<td>$\bar{\kappa}_L$</td>
<td>0.320</td>
<td>0.1091</td>
</tr>
<tr>
<td>$\bar{\kappa}_T$</td>
<td>0.066</td>
<td>0.0118</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.133</td>
<td>0.0454</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.232</td>
<td>0.0333</td>
</tr>
</tbody>
</table>

Mean log-likelihood: 2.753

### Table 7: Model Predictions

<table>
<thead>
<tr>
<th></th>
<th>Estimation Sample</th>
<th>Holdout Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Data$ $Model$ $\Delta \sigma_w$ $\Delta \bar{a}$</td>
<td>$Data$ $Model$ $\Delta \sigma_w$ $\Delta \bar{a}$</td>
</tr>
<tr>
<td>$Pr(U)$</td>
<td>11.05 30.09 30.85 30.09</td>
<td>60.28 61.38 60.58 61.38</td>
</tr>
<tr>
<td>$Pr(L)$</td>
<td>5.89 5.44 5.48 5.13</td>
<td>1.81 0.14 0.15 0.16</td>
</tr>
<tr>
<td>$Pr(C)$</td>
<td>39.40 32.17 9.04 32.00</td>
<td>0.13 5.68 18.75 5.61</td>
</tr>
<tr>
<td>$Pr(P)$</td>
<td>21.74 20.04 40.67 20.37</td>
<td>0.52 18.03 6.83 18.12</td>
</tr>
<tr>
<td>$E[\ell_C</td>
<td>C]$</td>
<td>3.11 4.81 3.37 4.81</td>
</tr>
<tr>
<td>$E[\ell_P</td>
<td>P]$</td>
<td>1.89 3.91 3.73 3.91</td>
</tr>
<tr>
<td>$E[\ell_T]$</td>
<td>1.11 1.99 2.01 2.03</td>
<td>0.01 0.08 0.07 0.08</td>
</tr>
</tbody>
</table>

Notes: $Data$ represents raw means from the respective sample, $Model$ the mean prediction, $\Delta \sigma_w$ the mean prediction when each borewell in the estimation (holdout) sample is assigned a $\sigma_w$ drawn at random from the holdout (estimation) sample, and $\Delta \bar{a}$ the mean prediction when each borewell in the estimation (holdout) sample is assigned an $\bar{a}$ drawn at random from the holdout (estimation) sample.
Appendix

A Proof of Proposition 2

We begin by proving that under any water-transfer arrangement economic surplus, \( V_j(r) \) \( j = U, A, C, P \), strictly diminishes with mean-preserving increases in uncertainty \( r \). But first, we paraphrase a result from Diamond and Stiglitz (1974, p. 340, fn. 8). For any twice continuously differentiable function \( h(w) \)

\[
\int h(w) \psi_r dw = \int h(w) T(w, r) dw, \tag{A.1}
\]

where \( \psi \) is a p.d.f., \( \psi_r \) is its partial derivative with respect to \( r \), and \( T(w, r) \) is a nonnegative function defined in equation (5) of Diamond and Stiglitz (1974).

Lemma 1 If \( g \) is strictly concave, then \( V'_j(r) < 0 \) for \( j = U, A, C, P \).

Proof. (i) \( V_U \): Differentiating definition 2 and using the envelope theorem yields

\[
V'_U(r) = \ell_U \left[ \int f(w/\ell_U) \psi_r dw - c \right]. \tag{A.2}
\]

Since, by (A.1) and the concavity of \( f \), the first term in square brackets is negative, \( V'_U(r) < 0 \).

(ii) \( V_A \): Follows from (A.2) with \( \ell_U \) replaced by \( a \).

(iii) \( V_C \): As noted in fn. 17, \( V_C = a E \left[ f\left(\frac{w-\tau_C}{a}\right) - c \right] + \ell_C \left[ f\left(\frac{\tau_C}{\ell_C}\right) - c \right] \). Differentiating with respect to \( r \) and using the envelope theorem leads to an expression analogous to (A.2).

(iv) \( V_P \): Given the discussion in fn. 17,

\[
V_P = a E \left[ f\left(\frac{w - \tau_P}{a}\right) - c \right] + \ell_P \left[ f\left(\frac{\tau_P}{\ell_P}\right) - c \right]
\]

\[
= (a + \ell_P) E \left[ f\left(\frac{w}{a + \ell_P}\right) - c \right], \tag{A.3}
\]

where the second line follows from the ex-post efficiency condition \( f'(\frac{\tau_P}{\ell_P}) = f'(\frac{w-\tau_P}{a}) \). Dif-
Differentiating with respect to \( r \) in this case yields
\[
V'_P(r) = E \left[ g \left( \frac{w}{a + \ell_P} \right) - c \right] \frac{\partial \ell_P}{\partial r} + (a + \ell_P) \left[ \int f \left( \frac{w}{a + \ell_P} \right) \psi_r dw - c \right] 
\]
(A.4)
\[
= c \left( \frac{1}{\eta} - 1 \right) \frac{\partial \ell_P}{\partial r} + (a + \ell_P) \left[ \int f \left( \frac{w}{a + \ell_P} \right) \psi_r dw - c \right],
\]
where the second line uses equation (7). In the case of the per-irrigation arrangement, a precise analog to Proposition 1 applies. Thus, given that \( g \) is concave, \( \partial \ell_P / \partial r < 0 \). The first term of (A.4) must, therefore, be negative and, using (A.1) again, the second term must also be negative.

**Proof of proposition 2:** (a) Proposition 1 implies that for sufficiently large \( r \), say \( r_U, \ell_U = a \) and, hence, \( V_U(r_U) = V_A(r_U) \). Recall that under perfect certainty, \( V_C(0) = V_U(0) > V_A(0) \). Given \( \tau_C(0) < w_L \) (see fn. 18) and lemma 1, \( V_C \) is continuously decreasing in \( r \) until it equals \( V_A \) at some \( r = r_C \). Now use equation (7) under perfect certainty to define \( \eta = c/g(w/a) \). For \( \eta \leq \eta, \ell_P = 0 \forall r \) and, consequently, \( V_P(r, \eta) = V_A(r) \forall r \) and, in particular, for \( r = 0 \). Define \( r_P(\eta) \) as the solution to
\[
V_P(r_P(\eta), \eta) = V_A(r_P(\eta)). \tag{A.5}
\]

Thus, clearly, \( r_P(\eta) = 0 \). Recall, also, that \( V_P(r, 1) = V_U(r) \) \( \forall r \), so \( r_P(1) = r_U \). To prove part (a), it is sufficient to show that \( r_P(\eta) \in (r_C, r_U) \) for some \( \eta \). This is so because \( V_P(0, \eta) < V_C(0) \) for all \( \eta < 1 \) and, therefore, by continuity, the functions \( V_C(r) \) and \( V_P(r, \eta) \) must cross at \( r^* < r_C \) if indeed \( r_P(\eta) > r_C \) (see figure 1).

Thus, it is sufficient to show that \( r'_P(\eta) > 0 \) so that as \( \eta \) is increased from \( \eta \) to \( 1 \) \( r_P(\eta) \) eventually exceeds \( r_C \). Differentiating equation (A.5) with respect to \( \eta \), substituting from equation (7), and rearranging gives
\[
c \left( \frac{1}{\eta} - 1 \right) \frac{\partial \ell_P}{\partial \eta} + \left[ \int h(w) \psi_r dw \right] r'_P(\eta) = 0 \tag{A.6}
\]
where \( h(w) = (a + \ell_P)f\left(\frac{w}{a + \ell_P}\right) - af\left(\frac{w}{a}\right) \). Since \( \partial \ell_P / \partial \eta < 0 \), and given (A.1), we have that \( \text{sign}(r'_P(\eta)) = -\text{sign}(h_{ww}(w)) \). Differentiating \( h(w) \) twice, we get
\[
h_{ww}(w) = \frac{1}{a + \ell_P} \left[ f''\left(\frac{w}{a + \ell_P}\right) - \frac{a + \ell_P}{a} f''\left(\frac{w}{a}\right) \right] \tag{A.7}
\]
A Taylor expansion around $\ell_P = 0$ gives

$$f''\left(\frac{w}{a+\ell_P}\right) \approx f''\left(\frac{w}{a}\right) - \frac{w\ell_P}{a^2} f'''\left(\frac{w}{a}\right).$$

Substituting into equation (A.7) and rearranging yields

$$h_{ww}(w) \approx \frac{-\ell_P}{a(a + \ell_P)} \left[ f''\left(\frac{w}{a}\right) + \frac{w}{a} f'''\left(\frac{w}{a}\right) \right]. \quad (A.8)$$

Since the term in square brackets is just $-g''(w/a)$, the concavity of $g$ ensures that $h_{ww} < 0$ and hence that $r'_P(\eta) > 0$.

(b) Having just established that $r_P(\eta) > r_C$ for some $\eta \in (\eta, 1)$, it must be true that $V_P(r, \eta) > V_C(r)$ over the interval $(r^*(\eta), r_P(\eta))$. ■

### B Additional Figures
Figure B.1: Rising Importance of Borewell Irrigation in India
Figure B.2: AVERAGE DEPTH TO WATER TABLE IN ANDHRA PRADESH
Figure B.3: Increasing Land Fragmentation in India