Introduction

Land tenancy contracts can be classified into two broad categories. (1) Owner cultivation where the owner cultivates the land himself possibly using hired labor. (2) A land lease agreement where the landowner leases out the land to be cultivated by a tenant and in turn asks for a fixed rental payment every period (a fixed rent contract) or a pre-determined share of the output (sharecropping contract). There is a large body of literature that addresses the issue of land tenure contracts [see Singh (1989), Basu (1996) and Ray (1998) for excellent surveys].

All three contractual forms often co-exist in close proximity. Shaban (1987) in a study of eight Indian villages finds that often even adjoining plots of land are characterized by different contractual forms. The question therefore is: what explains the actual choice of contract on a particular plot of land? This paper uses data from three villages in India to examine the question. We examine the effect of cultivating household and plot characteristics to examine the type of contract that is offered on a particular plot. Prior theoretical and empirical research in the area has treated the choice of a tenancy contract as one that involves choosing one out a set with three elements {wage, rent and share}. See Hallagan (1978), Allen (1982, 1985) and Eswaran and Kotwal (1985) as representative theoretical papers and Allen and Lueck (1992, 1999) as representative empirical papers. However as Bell (1995) points out "...any attempts to integrate land and labor contracts must come to grips with two salient features of agrarian organization in Asia, namely, that cultivating households make very extensive use of casual labor, and that most landowning households are not active in the market for tenancies." This essentially means that the decision for the owner whether to lease out the plot to be cultivated by a tenant or to cultivate the plot himself is quite distinct from the decision on what kind of a contract to offer, conditional on the decision to lease out the plot. We therefore think that the choice of a tenancy contract is better modelled as a sequential choice. In the sequential model the choice of a contract is a two step problem - the landowner first decides whether to cultivate the plot on his own or lease it out to be cultivated by a tenant. Once the landowner decides to lease out the plot, he has to decide whether to use a fixed rent or a share cropping contract. There is an alternative way of modelling contract choice where the landowner makes a simultaneous choice of one out of three available options {owner cultivation, fixed rent tenancy and share tenancy} to maximize his pay-off. This is the usual approach in the extant literature as mentioned above. For the sake of completeness and for purposes of comparison we derive estimates using a simultaneous choice model as well. Incidentally the co-existence of different contract forms is not specific to agricultural contracts. They are seen in business format franchering [see Lafontaine (1992)]. Hsiao et.al. (1998) report similar findings in Chinese Township and Village Enterprises.

Our estimation results show the following. (1) An increase in the age of the household head is associated with a significant increase in the probability that the household is offered a tenancy contract. (2) The greater the value of the plot, the greater is the probability that the plot is under owner cultivation. Second, conditional on the plot being under tenancy, the higher the value of the plot, the higher is the probability that the plot is under share tenancy. Given that the value of the plot is a measure of plot quality, we argue that the most productive plots are cultivated by the owner, the least productive plots are cultivated by a fixed rent tenant and the intermediate plots are cultivated by a share tenant.

The rest of the paper is organized as follows. Section 2 presents an overview of the data and selected descriptive statistics. Section 3 presents the the econometric methodology used. Section 3.1 discusses the sequential choice problem which is our principal focus in this paper. Section 3.2

analyses the simultaneous choice issue. Section 4 discusses the results. Finally Section 5 concludes.

Data

The data set used in this study comes from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and was part of ICRISAT's longitudinal Village level surveys in the semi-arid tropics of India. The survey was conducted over the period 1975-1984 but the labor market data exists only for 1979-1984. This is the period that we consider. We use data from the villages of Aurepalle, Shirapur and Kanzara which are situated in South-Central India (in the states of Andhra Pradesh and Maharashtra) and are predominantly agricultural with more than 94% of the households dependent on agriculture as the main source of income (either as cultivators or farm laborers). Even by Indian standards these villages are poor with a monthly per capita income of Rs. 700 (averaged over the survey period at 1977 prices) compared to the All-India per capita monthly income of Rs 1080 using the same base year. The data is a stratified sample of 40 randomly chosen households in each village, 10 in each of the four categories: (1) large farmers owning more than 3.2 acres in Aurepalle and more than 5.3 acres in Shirapur and Kanzara; (2) medium farmers owning between 1.2 and 3.2 acres in Aurepalle, between 2 and 5.3 acres in Shirapur and between 1.8 and 5.3 acres in Kanzara; (3) small farmers owning between 0.2 and 1.2 acres in Aurepalle, between 0.2 and 2.0 acres in Shirapur and between 0.2 and 1.8 acres in Kanzara; and finally (4) landless laborers who own less than 0.2 acres. The richness of the data from these surveys, both in terms of the breadth of information conveyed and the level of detail pertaining to each aspect of household decision making is amply illustrated by the numerous studies that have been conducted by economists around the world using this data set. [Walker & Ryan (1990) and Singh, Binswanger & Jodha (1985) provide details of the region and the survey]. Table 1 presents selected descriptive for some of the variables that are included in the regressions.

Our interest is in investigating the nature of the contract under which each plot is cultivated and to this end we examine data from each plot of land under consideration. We use data from 375 plots in Aurepalle, 256 plots in Shirapur and 604 in kanzara, giving us a total of 1235 plots. The majority of plots are under owner cultivation - in fact the data shows that 83% of all plots are under owner cultivation. Owner cultivation is particularly widespread in Aurepalle and Kanzara - 91.73% of the plots in Aurepalle and 85.43% of the plots in Kanzara are under owner cultivation. This is not particularly surprising because Singh (1988) argues that owner cultivation, using hired labor, is by far the most common cultivation form in Indian agriculture. In Shirapur share contracts are much more prevalent - nearly 35% of plots in Shirapur are under share tenancy. Note that only 23 of the 1235 plots are under fixed rent tenancy. [See Table 2 for a distribution of plots, classified on the basis of cultivation status].

Table 2 also presents a breakdown of the plots on the basis of soil characteristics and the main form of irrigation used on the plot. Notice that there is significant variation in the soil types across the three villages. For example in Kanzara, more than 85% of the plots have Medium Black soil, while in Shirapur majority of the plots have either Medium to Shallow Black or Gravelly soil. Finally in Aurepalle majority of the plots have Shallow Red soil. As far as irrigation is concerned, majority of the plots in all three villages are irrigated using a well with a traditional device.

Contract Choice: Econometric Methodology

We consider two alternative forms of contract choice. The first is the sequential choice model which is our principal interest in this paper. In this case, the landlord first decides whether to cultivate the plot on his own (possibly using hired labor) or to lease it out to be cultivated by a tenant. The landlord therefore chooses one contract out of the following two: $\{OWNER, LEASE\}$. If he decides to lease it out, then the question is whether to lease the plot out to be cultivated by a share tenant or a fixed rent tenant. footnote

As we mentioned before the existing theoretical literature on contract choice has almost exclusively focussed on a simultaneous choice model where the landlord, given a choice of three different contract forms, chooses the one that is likely to yield the maximum pay-off. In this case the landlord chooses one contract out of a set of three: $\{OWNER, RENTAL, SHARE\}$. We will examine both the sequential [see section 3.1] and the simultaneous choice [see section 3.2] problems.

Sequential Contract Choice

Contract choice is modelled as a two stage problem. In the first stage the owner has to decide whether to cultivate the plot himself or lease it out to be cultivated by a tenant. In the second stage, conditional on the owner choosing a tenancy contract, he has to decide whether to lease the plot out on a fixed rent or a share-cropping contract.

In the first stage define a variable *TENANT* such that

$$TENANT = \begin{cases} 0 \text{ if the plot is cultivated by the owner} \\ 1 \text{ if the plot is cultivated by a tenant} \end{cases}$$

Let us assume that there is an underlying response variable $TENANT^*$, which is unobservable and has a linear specification of the form

$$TENANT^* = \beta X + u, u \backsim IN(0, \sigma_u^2) \qquad \#$$

and X is a vector of agent and plot characteristics. While $TENANT^*$ is not observable what we do observe is a dummy variable *TENANT* such that

$$TENANT = \begin{cases} 0 \text{ if } TENANT^* \leq 0 \\ 1 \text{ otherwise} \end{cases}$$

The probability that $TENANT = 0$ (the land is cultivated by the owner himself) is $\Phi(-\beta X)$ while the probability that $TENANT = 1$ (the land is cultivated by a tenant) is $\Phi(\beta X)$ giving us the log likelihood function

$$L(\beta) = \sum_{TENANT=0} \ln \Phi(-\beta X) + \sum_{TENANT=1} \ln \Phi(\beta X)$$

Indexing plots by p and households (agents) by h, the estimating equation is

$$TENANT_{ph} = \beta_0 + \beta_1 X_{1h} + \beta_2 X_{2p} + u_{ph}; p = 1, ..., P; h = 1, ..., H \qquad \#$$

here X_{1h} represents agent characteristics and X_{2p} represents plot characteristics.

There is a second stage to the problem. Once the landlord has decided to lease out the land he has to decide whether to choose a fixed rent contract or a sharecropping contract. Define a variable *TENANT*2, such that

$$TENANT2 = \begin{cases} 0 \text{ if the plot is under fixed rent} \\ 1 \text{ if the plot is under share-cropping} \end{cases}$$

Let us assume that the underlying response variable for *TENANT*2 is $TENANT2^*$ which has the form

$$TENANT2^* = \gamma Z + e, e \backsim IN(0, \sigma_e^2) \quad \#$$

As in the first stage of the problem the estimating equation in the second stage can be written as

$$TENANT2_{ph} = \gamma_0 + \gamma_1 Z_{1h} + \gamma_2 Z_{2p} + e_{ph}; p = 1, ..., P; h = 1, ..., H \quad \#$$

where Z_{1h} represent agent (household) characteristics and Z_{2p} represent plot characteristics.

However $TENANT2$ is observed only when $TENANT = 1$, and so at the second stage we have a censored sample. Since $TENANT2$ is not a continuous variable, Heckman's two step procedure to correct for sample selection will not lead to consistent estimates. Equations (ref: eq1) and (ref: eq3) together constitute a bivariate qualitative dependent variable model that is characterized by partial observability. The set-up is as follows:

$$TENANT^* = \beta X + u; TENANT = \begin{cases} 0 \text{ if the plot is cultivated by the owner} \\ 1 \text{ if the plot is cultivated by a tenant} \end{cases}$$

$$TENANT2^* = \gamma Z + e; TENANT2 = \begin{cases} 0 \text{ if the plot is under fixed rent cultivation} \\ 1 \text{ if the plot is under share cropping} \end{cases}$$

and $TENANT2$ is observed only when $TENANT = 1$.

The model is one of partial observability, because we only observe three possible outcomes

1. $TENANT = 0$
2. $TENANT = 1, TENANT2 = 0$
3. $TENANT = 1, TENANT2 = 1$

The corresponding log likelihood function for the sample of P plots and H households can be written as [see Meng & Schmidt (1995)]

$$L(\beta, \gamma, \rho) = \sum_h \sum_p \begin{bmatrix} TENANT_{ph} TENANT2_{ph} \ln F(\beta X_{ph}, \gamma Z_{ph}; \rho) + \\ TENANT_{ph}(1 - TENANT2_{ph}) \ln[\Phi(\beta X_{ph} - F(\beta X_{ph}, \gamma Z_{ph}; \rho))] + \\ (1 - TENANT_{ph} \ln(1 - \Phi \beta X_{ph})) \end{bmatrix} \quad \#$$

where $F(.)$ and $\Phi(.)$ denote the bivariate standard normal cumulative distribution function and the univariate standard normal cumulative distribution respectively. Estimates of the parameters are obtained by maximizing the log likelihood function in equation (ref: ln2). The joint approach accounts for the potential correlation between the two error terms and corrects for potential bias in sample selection that would be incurred by estimating equations (ref: eq1) and (ref: eq3) separately.

Note that if the estimated correlation coefficient (ρ_{ue}) between the two-error terms u and e is not significant, it implies that the error terms in equations (ref: eq1) and (ref: eq3) are not correlated and we can estimate the second stage equation (estimate $TENANT2$ (equation (ref: eq4)) separately by running a binary probit in stage 2.

The vector of explanatory variables consists of cultivating household and plot characteristics. Household characteristics include the following: age, square of the age, sex and marital status of the household-head (*AGEHEAD*, *AGE*2*HEAD*, *SEX* and *MARITAL*1 respectively), dummies for the disability status of the adult members in the household (*MALEILL* and *FEMILL*). footnote The regressions also control for the composition of each household - including total number of males, females, children and old persons (*TOTMAL*, *TOTFEM*, *TOTCHILD* and *TOTOLD* respectively). The age of the household head (*AGEHEAD*) and the square of the age (*AGE*2*HEAD*) are proxies for the level of experience of the cultivator. See Table A1 for a descritopn of all the explanatory variables used in our analysis.

Plot characteristics include value of the plot in rupees (*VALUE*), dummies for the main source of irrigation for the plot footnote , and for alternative soil types footnote and the percentage of total cultivated area that is irrigated (*IRR*). *VALUE* is the monetary value of a plot and is a proxy for land quality. Per acre estimated value of the plot in Rs. 100 were recorded based on information obtained from either the patwari (land assessor)or some other knowledgeable person in the village. While recording the value of the plot, potential sale value of the plot, location of the plot, irrigation and topography are taken into account.

Finally we include two village dummies $(DV1 \text{ and } DV2)$ to account for any unobserved heterogeneity. $DV1 = 1$ if Aurepalle and $DV2 = 1$ if Shirapur. The reference case is that the village is Kanzara.

We would like to point out that the data set is indeed a panel. But we do not include the panel aspect of the data because the average length of a contract in the ICRISAT region is more than 14 years and one might assume without the loss of generality that the contracts are stable for each plot over the 6 years that we consider.

Simultaneous Contract Choice

In this case the principal chooses one of the three contracts {Owner Cultivation, Fixed Rent, Share} to maximize his profits. This is a problem with a polychotomous dependent variable, which can take three distinct values. A variety of qualitative response models have been devised to deal with such cases. They fall into two types - models designed to deal with ordered responses and models designed to deal with unordered responses. Since *ex ante* there is no obvious way to order the three contracts, we believe that a model designed to deal with unordered responses is the appropriate one to use in this context. Hence we derive our parameter estimates using a multinomial logit model.

Define a variable *CONTRACT* such that

$$CONTRACT = \begin{cases} 0 \text{ if the plot is under owner cultivation} \\ 1 \text{ if the plot is under fixed rent tenancy} \\ 2 \text{ if the plot is under share-cropping} \end{cases}$$

We wish to examine the choice of contract for each plot. As before, assume that there is an underlying response variable $CONTRACT^*$ defined by the following:

$$CONTRACT^* = \beta X + \varepsilon; \varepsilon \backsim IN(0, \sigma_\varepsilon^2)$$

Now $CONTRACT^*$ is unobservable and instead we observe the variable *CONTRACT* such that

$$CONTRACT = \begin{cases} 0 \text{ if } CONTRACT^* \leq \mu_1 \\ 1 \text{ if } \mu_1 \leq CONTRACT^* \leq \mu_2 \\ 2 \text{ if } CONTRACT^* \geq \mu_2 \end{cases}$$

We normalize $\mu_1 = 0$. Then the probability that the plot is under owner cultivation is given by

$\Pr\{CONTRACT = 0\}$ which is equal to $\Phi(-\beta X)$; the probability that the plot is under fixed rent is given by $\Pr\{CONTRACT = 1\}$ which is equal to $\{\Phi(\beta X - \mu_2) - \Phi(-\beta X)\}$and finally the probability that the plot is share-cropped is given by $\Pr\{CONTRACT = 2\}$ which is $\Phi(\beta X - \mu_2)$. Then the log-likelihood function can be written as

$$L(\beta;\mu_2) = \sum_{CONTRACT=0} \log(\Phi(-\beta X)) + \sum_{CONTRACT=2} \log(\Phi(\beta X - \mu_2)) + \sum_{CONTRACT=1} \log(\Phi(\beta X - \mu_2) - \log(\Phi(-\beta X)))$$

The vector of explanatory variables are exactly the same as before and hence we do not discuss them again.

Results

Let us now turn to the estimation results. We first present the results from the sequential choice model [Section 4.1] and then the results from the simultaneous choice model [Section 4.2].

Results from the Sequential Choice Model

The maximum likelihood probit results for the first stage estimation (choice of cultivation status) are presented in Table 3. A positive value of the coefficient indicates that the variable increases the probability of tenant cultivation while a negative value of a coefficient indicates that the variable decreases the probability of tenant cultivation. Relative to Kanzara, the probability that a plot is under tenant cultivation is significantly higher in Shirapur. The model has substantial predictive power – in particular, the model predicts correctly more than 86% of the time.

Let us first examine the cultivating household characteristics. Notice first that *AGEHEAD* is positive and significant. This implies that the higher the age of the head of the cultivating household, the greater is the probability that the plot is cultivated by a tenant, indicating that holding everything else constant, older agents are more likely to be employed as tenants, while younger agents are more likely to be cultivating the plots of land on their own (possibly with the help of wage labour). footnote This is not a surprising result since there exists a fairly old literature that argues that individuals start off being wage cultivators and graduate to becoming tenants [see Spillman (1919)]. Further, owners are also more willing to lease out land to older individuals because this way more information is available to the landlords regarding the type of the individual, in the presence of private information regarding individual characteristics. An increase in the total number of males in the household or an increase in the number of elderly members in the household reduce the probability that the plot is cultivated by a tenant, but an increase in the number of children in the household increases the probability that the plot is cultivated by a tenant.

Let us now turn to the plot characteristics. The most interesting result concerns the sign and significance of *VALUE*. Remember that *VALUE* is a measure of the land quality of a particular plot. We find that the coefficient of *VALUE* is negative and significant which implies that an increase in plot quality reduces the probability that the plot is cultivated by a tenant. The marginal results (not presented but available on request) show that a Rs 100000 increase in the value of the plot decreases the probability that the plot is under tenant cultivation by 37 percentage points. Of the other plot characteristics, all three irrigation dummies are negative and significant indicating that a plot where the main source of irrigation is a well with traditional device or a well with electric motor or a well with oil engine is less likely to be cultivated by a tenant. A plot where the soil is gravelly is less likely to be cultivated by a tenant as is a plot with a higher proportion of the area under irrigation. One could argue that the value of a plot is determined (at least partly) by the soil type of that plot, the kind of irrigation facility that is available on the plot and the proportion of the plot that is irrigated and hence it is meaningless to include both the value of the plot and the other plot level characteristics as explanatory variables. Therefore we examined an alternative specification where we included only *VALUE* and excluded all the other plot level characteristics. We do not report the results, because the results remain the same - *VALUE* is still negative and significant. In fact the marginal results show that in this case *VALUE* has a stronger negative effect on the probability of the plot being under owner cultivation. In this case the marginal results show that a Rs 100000 increase in the value of the plot reduces the probability that the plot is under tenant cultivation by more than 42 percentage points. Therefore while there is some incidental association between the value of the plot (as measured by *VALUE*) and other plot level characteristics, this association is not very strong.

Table 4 presents the estimated coefficients for the second stage - conditional on the plot being cultivated by a tenant, what are the factors that affect the probability that the plot is cultivated by a fixed rent tenant as opposed to a share tenant. Here we present two sets of results - the binary probit results for *TENANT*2 where we do not correct for sample selection and the Full Information Maximum Likelihood (FIML) results from the joint estimation of *TENANT* and *TENANT*2, where we correct for sample selection - the Partial Observability Model. Note that in the second stage the estimation sample is quite small and hence we are unable to include the entire set of explanatory variables as in stage 1 (Table 3). Note also that the correlation coefficient between the two error terms u and e is not significant. Therefore a binomial probit is a good benchmark in this situation and further we look at the results from the FIML estimates, we see that none of the explanatory variables significantly affect the conditional probability of the plot being cultivated by a share tenant. The binary probit results show that the conditional probability of the plot being under share tenancy is significantly lower in Aurepalle (relative to Kanzara). Age of the household head (*AGEHEAD*) is negative and significant indicating that the conditional probability of being a share tenant is significantly lower for an older tenant. However notice that there is a non-linearity in the effect of the age of the head of the household as indicated in the positive and significant sign of the square of the age of the household head (*AGE*2*HEAD*). We would like to point out that this finding about the age of the household head accords well with the ”agricultural ladder” hypothesis which states that the youngest workers will be wage-labourers, the intermediate ones share-croppers and the olders workers fixed-rent tenants. None of the other household level characteristics and none of the plot level characteristics excepting *VALUE* have a significant effect on the conditional probability that the plot is under share tenancy. *VALUE* is positive and significant - implying that an increase in the value of the plot increases the conditional probability that the plot is under share tenancy. As in Stage 1 we re-estimate the model ignoring all the plot level characteristics with the exception of *VALUE*, since it could be argued that the value of a particular plot is at least partly determined by the other plot level characteristics. As before the age of the head (*AGEHEAD*) of the cultivating household has significant and negative effect on the conditional probability that the plot is under share tenancy. There is again a significant non-linearity in the age of the household head effect - *AGE*2*HEAD* is positive and significant. In this case however the value of the plot (*VALUE*) has a significant and positive effect on the conditional probability that the plot is under share tenancy. footnote

Notice that *VALUE* is positive and significant in step 2 [see Table 4]. It implies that as the value of a plot increases, the conditional probability that it is cultivated by a share tenant increases. This result, combined with the fact that *VALUE* is positive and significant in stage 1 [Table 3] implies the following. First, the highest quality plots are cultivated by the owner. Second, conditional on the plot being leased out, the higher quality plots are cultivated by a share tenant. So essentially the highest quality plots are cultivated by the owner, the lowest quality plots are cultivated by a fixed rent tenant and the medium quality plots are cultivated by a share tenant.

We re-estimated the model with only household characteristics as the set of explanatory variables (ignoring the plot level characteristics) and only plot characteristics (ignoring the household level characteristics) as the set of explanatory variables. In both cases however we do include the village dummies. The predictive power of the model is slightly reduced in both cases, though not significantly. The estimated coefficients are presented in Tables 5 and 6 respectively. In neither of the cases is the null hypothesis of zero correlation between the two error terms u and e (ρ_{ue}) rejected. So in both cases, binary probit in Stage 2 gives consistent estimates. Notice from Table 5, in the Stage 1 regressions, none of the cultivating household characteristics, with the exception of *TOTMAL* have a significant effect on the probability of the plot being under tenancy. Turning to the Stage 2 estimates, we find that the age of the household head and the square of the age of the household head are both significant and have opposite signs. Once again we find that as the age of the cultivating household head increases the conditional probability that the worker is offered a tenancy contract goes up. When we re-estimate the model with only plot characteristics, the results remain unaffected. In particular, *VALUE* is negative and significant in stage 1 and is positive and significant in stage 2 - see Table 6. This implies that as before better quality plots are owner cultivated but conditional on the plot being under tenancy, better quality plots are under fixed rent tenancy.

Finally we added two dummies *LMWD* and *LFWD* as additional household characteristics. *LMWD* equals one if any adult male member of the household worked in the village labour market in the previous year and *LFWD* equals one if any adult female member of the household worked in the village labour market in the previous year. Because of space constrainst, these estimation results are not presented. They are however available on request. In the first stage regressions, both *LMWD* and *LFWD* are negative and significant implying that the probability that the household being a tenant is significantly reduced by labour market participation of the members. *LMWD* and *LFWD* do not have a significant effect on the conditional probability of the plot being under share tenancy. footnote

Results for the Simultaneous Choice Model

The results for the simultaneous choice model are presented in Table 7. We only present the multinomial logit estimates for the complete case i.e., the set of explanatory variables includes both household and plot characteristics. We conduct similar robustness tests as in the sequential contract choice model and the results can be made available on request. The baseline category is the plot is under owner cultivation ($CONTRACT = 0$). The results are as expected. Relative to the baseline category, households with more children are more likely to be hired as tenants (both fixed rent and share) and households with more elderly are more likely to be hired as share tenants. A plot where the soil is Shallow Red is less likely to be leased out to a tenant while a plot where the soil is gravelly is less likely to be leased out to a share tenant. If we examine the signs of *VALUE* we find that an increase in the value of the plot moves the plot away from both fixed rent cultivation and share cropping to owner cultivation. The marginal results show that a Rs 100000 increase in the value of the plot increases the probability of owner cultivation by 29 percentage points and reduces the probability of fixed rent tenancy by 1 percentage point and reduces the probability of share tenancy by 28 percentage points. This is interesting and is further proof of the oft repeated argument that owners tend to cultivate the best quality plots themselves.

Conclusion

While there has been quite a bit of work on the Marshallian inefficiency of sharecropping contracts, there has been a significant void in empirical work analyzing factors affecting the choice of contracts. This paper seeks to fill that gap. Our estimation results show the following. An increase in the age of the household head is associated with a significant increase in the probability that the household is offered a tenancy contract. Turning to plot characteristics we find an interesting relationship between the value of a particular plot and the contract offered. First, the greater the value of the plot, the greater is the probability that the plot is under owner cultivation. Second, conditional on the plot being under tenancy, the higher the value of the plot, the higher is the probability that the plot is under share tenancy. The value of the plot is a measure of plot quality. We therefore argue that the most productive plots are cultivated by the owner, the less valuable plots are cultivated by a fixed rent tenant and the intermediate plots are cultivated by a share tenant.

References

1. Allen, D. W. & Lueck, D. (1992), 'Contract Choice in Modern Agriculture: Cash Rent versus Cropshare', *Journal of Law and Economics*, **XXXV**, 397–426.
2. Allen, D.W. & Lueck, D. (1999), 'The Role of Risk in Contract Choice', *Journal of Law, Economics and Organization*, 15(3), pp. 704-36.
3. Allen, F. (1982), 'On Share Contracts and Screening', *Bell Journal of Economics*, 13(2), pp. 541-7
4. Allen, F. (1985), 'On the Fixed Nature of Sharecropping Contracts', *Economic Journal*, 95 (377), pp. 30-48
5. Alston, L. J., Dutta, S. K. & Nugent, J. B. (1984), 'Tenancy Choice in a Competitive Framework with Transactions Costs', *Journal of Political Economy*, **92**(6), 1121–1133.
6. Basu, K. (1984), *The Less Developed Economy*, Oxford University Press.
7. Bell, C. (1995), Review of *The Economics of Contract Choice*: *An Agrarian Perspective* by Yujiro Hayami and Keijiro Otsuka, *Journal of Economic Literature*, 33(1), pp. 270-1.
8. Eswaran, M. & Kotwal, A. (1985), 'A Theory of Contractual Structure in Agriculture', *American Economic Review*, **75**(3), 352–367.
9. Hallagan, W. (1978), 'Self Selection by Contractual Choice and the Theory of Sharecropping', *Bell Journal of Economics*, 9, pp. 344-54
10. Hsiao, C., Perrigne, I, Nugent, J and Qiu, J. (1998), 'Shares versus Residual Claimant Contracts: The Case of Chinese TVEs', *Journal of Comparative Economics*, **26**(2), 317-37.
11. Lafontaine, Francine, (1992), 'Agency Theory and Franchering: Some Empirical Results', *Rand Journal of Economics*, v 23, pp. 263-83
12. Meng, C. & Schmidt, P. (1985), 'On the Cost of Partial Observability in the Bivariate Probit Model', *International Economic Review*, **26**(1), 71–85.
13. Shaban, R. A. (1987), 'Testing Between Competing Models of Sharecropping', *Journal of Political Economy*, **95** (5), 893–920.
14. Singh, N. (1989), "Theories of Sharecropping", in P. Bardhan (ed.) *The Economic Theory of Agrarian Institutions*, Oxford University Press.
15. Singh, R. P., Binswanger, H. P. & Jodha, N. S. (1985), *Manual of Instructions For Economic Investigators in ICRISAT's Village Level Studies*, Technical report, ICRISAT, Patancheru, Andhra Pradesh 502324, India.

16. Walker, T. S. & Ryan, J. G. (1990), *Village and Household Economies in India's Semi - Arid Tropics*, The Johns Hopkins University Press, Baltimore and London.

Table 1: **Selected Descriptive Statistics for Cultivating Household Characteristics**

	Entire Sample			**Plots under Tenancy**		
Variable	**N**	**Mean**	**Count**	**N**	**Mean**	**Count**
AGE	1235	48.66		210	49.71	
TOTMAL	1235	2.13		210	1.94	
TOTFEM	1235	1.75		210	1.81	
TOTCHILD	1235	2.97		210	2.86	
TOTOLD	1235	0.44		210	0.23	
			67[b]			9[b]
MARITAL	1235		131[a]	210		15[a]
			1104[b]			195[b]
MALEILL	1235		1066[a]	210		187[a]
			169[b]			23[b]
FEMILL	1235		1224[a]	210		209[a]
			11[b]			1[b]

Notes:
[a] : Count = 0
[b] : Count = 1

Table 2: **Number of Plots**, **Classified by Ownership Status**, **Soil Type and Irrigation Source**

	Aurepalle	**Shirapur**	**Kanzara**
Ownership Status			
Owner Operated	344	165	516
Fixed Rent	13	0	10
Share Tenancy	18	91	78
Total under Tenancy	31	91	88
Soil Type			
Deep Back	0	19	57
Medium Black	48	39	518
Medium to Shallow Black	63	102	29
Shallow Red	253	0	0
Gravelly	3	91	0
Problem Soil*	7	1	0
Others	1	4	0
Irrigation Source			
Well with Traditional Device	291	221	588
Tank	2	0	1
Canal	2	0	0
Well with Electric Motor	80	15	15
Well with Oil Engine	0	19	0
Other	0	1	0

Notes:

*: For Example Saline Soil

Table 3: **Binary Probit Estimates for First Stage Cultivation Status**

Dependent Variable *TENANT*

Variable	**Coefficient**	**Standard Error**
CONSTANT	-2.008	1.143
DV2	1.108**	0.190
DV1	-0.305	0.225
AGEHEAD	0.085**	0.035
AGE2HEAD	0.000	0.000
MARITAL1	0.603**	0.205
TOTMAL	-0.216*	0.062
TOTFEM	0.066	0.062
TOTCHILD	0.088***	0.032
TOTOLD	-0.756***	0.108
MALEILL	-0.261	0.176
FEMILL	-0.410	0.558
IRRD0	-1.855**	0.702
IRRD3	-1.961**	0.630
IRRD4	-2.441***	0.754
SOILD1	-0.134	0.220
SOILD3	0.147	0.171
SOILD5	-0.372	0.249
SOILD6	-0.916***	0.251
IRR	-0.010*	0.005
VALUE	-1.933E-04***	5.958E-05

Notes:

*: Significant at 10%

**: Significant at 5%

***: Significant at 1%

Table 4: **Binary Probit and FIML Estimates for Second Stage Tenancy Status Dependent Variable** *TENANT*2

Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
	Binary Probit Estimates		**FIML Estimates**	
CONSTANT	115.090*	61.770	74.413	154.357
DV2	7.324	170995.960		
DV1	-6.508**	2.629		
AGEHEAD	-4.532**	2.297	-2.804	5.977
AGE2HEAD	0.043*	0.022	0.026	0.056
TOTMAL	0.320	0.672	-0.643	2.387
TOTFEM	0.375	0.798	0.292	1.257
TOTCHILD	0.233	0.545	0.574	1.987
TOTOLD	0.514	2.169	0.561	5.287
MALEILL	3.937	3.604	4.438	12.129
FEMILL	2.571	3197910.200	2.861	201456.020
IRRD0	2.122	9.248	2.315	7.406
SOILD3	0.660	2.111	-1.994	5.700
SOILD5	0.425	4.216	-6.193	19.792
VALUE	0.002*	0.001	0.002	0.004
ρ_{ue}			-0.721	3.007

Notes:

*: Significant at 10%

**: Significant at 5%

***: Significant at 1%

Table 5: **Binary Probit Estimates for First and Second Stage**:

Cultivating Household Characteristics Only

Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
	First Stage Estimates		**Second Stage Estimates**	
CONSTANT	-4.217***	0.892	203.983***	73.157
DV2	0.610***	0.121	6.881	141572.440
DV1	-0.640***	0.133	-7.777***	2.969
AGEHEAD	0.099***	0.036	-7.714***	2.778
AGE2HEAD	-0.001*	3.439E-04	0.073***	0.026
MARITAL1	0.530***	0.197		
TOTMAL	-0.267***	0.058	-0.105	0.494
TOTFEM	0.056	0.061	0.566	0.637
TOTCHILD	0.079**	0.031	0.241	0.389
TOTOLD	-0.680***	0.101	0.328	1.677
MALEILL	-0.184	0.161	5.865**	2.730
FEMILL	-0.599	0.557	1.843	1947571.000

Notes

*: Significant at 10%

**: Significant at 5%

***: Significant at 1%

Table 6: **Binary Probit Estimates for First and Second Stage**:

Plot Characteristics Only

Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
	First Stage Estimates		**Second Stage Estimates**	
CONSTANT	1.204*	0.714	0.074	1.051
DV2	0.992***	0.170	6.663	172377.740
DV1	-0.173	0.203	-1.172*	0.695
IRRD0	-1.938***	0.698	0.277	0.978
IRRD3	-2.024***	0.621		
IRRD4	-2.133***	0.745		
SOILD1	0.070	0.210		
SOILD3	0.246	0.163	0.150	0.636
SOILD5	-0.184	0.231	0.133	0.762
SOILD6	-0.750***	0.226		
IRR	-0.009**	0.005		
VALUE	-1.958E-04***	5.604E-05	6.818E-04**	3.250E-04

Notes

*: Significant at 10%

**: Significant at 5%

***: Significant at 1%

Table 7: **Multinomial Logit Estimates for the Sequential Choice Model**

Dependent Variable: *CONTRACT*

Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
	Fixed Rent Contract		**Share Cropping Contract**	
CONSTANT	-152.458	3727998.000	-2.479	2.405
DV2	-41.879	2728062.400	2.071***	0.352
DV1	3.547**	1.467	-0.752	0.484
AGEHEAD	3.700	2.365	0.140*	0.076
AGE2	-0.031	0.022	-0.001	0.001
MARITAL1	34.281	3199289.600	0.886**	0.394
TOTMAL	-1.098*	0.585	-0.418***	0.123
TOTFEM	-0.088	0.512	0.102	0.121
TOTCHILD	0.927***	0.297	0.149**	0.062
TOTOLD	-6.150***	2.017	-1.187***	0.202
MALEILL	-0.332	0.968	-0.499	0.359
FEMILL	-22.927	10382971.000	-0.698	1.102
IRRD0	13.583	1913865.100	-3.777**	1.468
IRRD3	-40.347	3980029.800	-3.614***	1.324
IRRD4	5.668	9976413.700	-4.062***	1.558
SOILD1	-32.524	4152479.300	-0.132	0.409
SOILD3	-1.662	1.254	0.148	0.326
SOILD5	-4.881***	1.747	-0.962*	0.567
SOILD6	-25.959	3151152.800	-1.525***	0.459
IRR	0.230	19138.651	-0.026**	0.011
VALUE	-0.005***	0.001	-3.490E-04***	1.307E-04

Notes

*: Significant at 10%

**: Significant at 5%

***: Significant at 1%

Baseline Category: Plot is under Owner Cultivation

Table A1: Definition of Variables

Variable	Definition
DV1	=1, if village is Kanzara
DV2	=1, if village is Shirapur
MARITAL1	=1, if the household head is married
MALEILL	=1, if any adult male member of the household is disabled
FEMILL	=1, if any adult female member of the household is disabled
AGEHEAD	Age of the household head
AGE2HEAD	Square of the Age of the household head
TOTMAL	Total number of adult (working age) males in the household
TOTFEM	Total number of adult (working age) females in the household
TOTCHILD	Total number of children in the household
TOTOLD	Total number of elderly in the household
VALUE	Value of the plot in Rs ’000
IRR	Proportion of the plot that is irrigated
IRRD0	=1, if source of irrigation is well with traditional device
IRRD3	=1, if source of irrigation is well with electric motor
IRRD4	=1, if source of irrigation is well with oil engine
SOILD1	=1, if soil type is medium black
SOILD2	=1, if soil type is deep black
SOILD3	=1, if soil type is medium to shallow red
SOILD5	=1, if soil type is shallow red
SOILD6	=1, if soil type is gravelly