# Evidence on the Relationship between Risk and Incentives

Luis H. B. Braido*†

January 12, 2009

**Abstract**

The standard moral hazard theory predicts that, for each given agent, the pay-for-performance sensitivity of a linear compensation scheme should decrease with the project's underlying risk. The empirical literature has used data on franchising, executive compensation, and tenancy contracts—typically undertaken by different agents—to assess this relationship. These results could be biased since the optimal incentive power also depends on the agent's risk aversion. This paper uses data on farmers who manage multiple plots under different contracts to study the risk properties of different contracts designed for the same agent. The results do not support the risk-sharing prediction.

Keywords: Risk, incentives, principal, agent, tenancy, data. JEL Classification: C52, D82, O12, Q15.

*I am thankful for comments from Marcelo Fernandes, Daniel Ferreira, Helena S. Perrone, and seminar participants at the University of Illinois at Urbana-Champaign and the 2006 SBE Meeting.

†Graduate School of Economics, Getulio Vargas Foundation; Praia de Botafogo 190, s.1100, Rio de Janeiro, RJ 22253-900, Brazil; lbraido@fgv.br.

# 1 Introduction

In many environments, the principal is not able to monitor all decisions made by the agent, and therefore introduces payments based on performance to induce appropriate decisions. Contracts based on pay for performance create incentives at the cost of imposing risk on the agent. The risk faced by the agent increases with the marginal pay-for-performance sensitivity—also referred to as the contract's incentive power—and with the risk of the underlying project. Since risk-averse agents require higher average payments to accept riskier contracts, one would expect to observe a negative correlation between the power of contracts designed for each agent and the exogenous risk of the underlying project.

This intuitive idea holds in environments in which the principal is restricted to use linear contracts based on individual performance. When nonlinear schedules are considered, the inverse relation between risk and incentives depends on strong assumptions about preferences and shocks, as was formalized by Holmstrom and Milgrom (1987).[1] In practice, however, there are many situations in which one is not able to implement nonlinear contracts (or contracts based on relative performance) because tenants could transfer output from one plot to another.

In spite of the theoretical dispute on the exact predictions of moral hazard models, the risk-incentive tradeoff has motivated a large body of empirical works (which are discussed in Section 2). A major difficulty in the empirical analysis is that each agent is usually associated to a single contract. Hence, in cases studied in the literature, different contracts are associated with different risk environments and also with different agents. If preferences are heterogeneous, this type of data would not be appropriated to test the risk-sharing prediction, because it is possible that less risk-averse agents—who are willing to accept compensations schemes with high pay-for-performance sensitivity—are matched with riskier projects. This point was stressed by Ackerberg and Botticini (2002).

This paper uses a well-known data set on tenancy contracts, collected by the International Crops Research Institute for Semi-Arid Tropics (ICRISAT) in India,

[1] These authors study the problem of a risk-neutral principal designing a contract to a risk-averse agent with constant absolute risk aversion, $\rho$, and quadratic effort disutility, $\frac{k}{2}e^2$. In their model, the final revenue depends linearly on the agent's hidden effort, $e$, and on a normally distributed shock with zero mean and constant variance, $\sigma^2$. The solution for the continuous-time version of this problem, where output follows a Brownian motion and the agent controls the drift of such a process, is shown to be as if the problem were the static one and the principal were constrained to use linear contracts. Moreover, the incentive power of the optimal linear contract in the static setup is $\left(1 + k\rho\sigma^2\right)^{-1}$. Therefore, for each given agent (i.e., for fixed values of $k$ and $\rho$), there is an inverse relation between the incentive power of the optimal contract and the exogenous risk, $\sigma^2$.

which contains a significant fraction of households that simultaneously manage multiple plots under different land contracts. This feature of data allows one to compare contracts designed for the same agent, avoiding concerns about preference heterogeneity.

There are three general contract types in the data, namely: ownership, in which the farm is managed by its owner; fixed rent, in which the tenant pays an upfront fee for using the land, bears all input costs, and retains the entire revenue; and sharecropping, in which the landlord supplies the land, the tenant bears most input costs, and they share the final output.[2] The pay-for-performance sensitivity is maximal in lands managed under ownership and fixed rent, as the farmer retains 100% of the final output. This paper tests whether households managing multiple plots use the sharecropping contract in their riskier lands, as predicted by the risk-sharing argument.

The empirical strategy is divided in two steps. First, an econometric model is used to measure the plot's underlying risk. Lafontaine and Bhattacharyya (1995) and Allen and Lueck (1999) identify the exogenous risk under the assumption that the performance measure is composed of two independent and additive parts, one endogenous and another exogenous. In agriculture, however, the final output is typically determined by the interaction of endogenous input choices and exogenous climatic shocks. Since detailed information on cropping activities is available, one can use a stochastic Cobb-Douglas production function to model the interaction between the endogenous and exogenous variables affecting the final output.

Once a plot-specific measure of risk is constructed (from the estimated residuals of the production function), the Glesjer's heteroscedasticity test is adapted to assess how risk differs across plots under different contracts. Household fixed effects are used to compare the risk-incentive relationship in the plots cropped by the same agent. Household fixed effects also account for other agent-specific determinants of the contract design (such as credit constraints). The results do not support the risk-sharing prediction.

The remainder of the paper is organized in the following manner. The next section presents a brief review of the related empirical literature. Section 3 describes the data; Section 4 introduces the econometric setup; and Section 5 presents the results. Robustness of the results is checked in Section 6, and a brief conclusion appears in Section 7.

[2] These contracts are linear and depend only on the performance of each plot—i.e., no formal payment is conditional on relative performance. In many cases, the landlord shares the cost of some inputs and provides supervision.

## 2 Related Literature

Data on executive compensation, franchising, and land tenancy are commonly used in empirical papers about the incentive design of contracts. Lambert and Larker (1987); Garen (1994); Aggarwal and Samwick (1999); Jin (2002), and Dee, Lulseged, and Nowlin (2005) use data on executive compensation and find evidence supporting the existence of a trade-off between risk and incentives.[3] The literature on franchising finds distinct signals for the risk-incentive relationship. For instance, Lafontaine (1992) show that: (i) the royalty rate paid by the franchisee does not increase with the risk of discontinuation; but (ii) the fraction of franchised outlets is increasing in the average fraction of discontinued outlets in the sector. Moreover, Lafontaine and Bhattacharyya (1995) show that risk sharing is not necessary to explain the evidence on franchising contracts.[4]

This debate is also active in the tenancy literature. Rao (1971) argues that, in India, share contracts are extensive for products (such as rice) and areas (such as the northern regions) where risk and entrepreneurial profit are low. On the other hand, fixed-rent contracts are common in situations of high risk and significant scope for decision making (e.g., tobacco farms). Allen and Lueck (1992) study tenancy contracts from Nebraska and South Dakota in the U.S. and report that sharecropping contracts are more likely to be used in lands where the costs of dividing the output is low. They find that sharecropping contracts are associated with crops such as corn (high volatility) and wheat (low volatility), which suggests that there is no clear relation between risk and incentive power. Similarly, Allen and Lueck (1999) use four different land-leasing surveys (from Nebraska, South Dakota, British Columbia, and Louisiana) to test the relationship between risk and incentive power. They have data on the form of the tenancy contract (fixed rent or sharecropping) and the fraction of crop retained by the sharecropping tenant. Exogenous risk is measured by the variability of crop yield across plots of a certain region. Using a logit model, they show that risk does not have a negative impact on the probability of a crop being leased under fixed rent. Tobit regressions also indicate that share rates are not decreasing in risk.[5]

All these papers do not control for potential heterogeneity in agents' preferences. Ackerberg and Botticini (2002) stress that the endogenous matching between principals and agents could change the interpretation of previous results. Executives, franchisees, and tenants under high-powered contracts could be those with higher

[3] The results are only weakly significant in Garen (1994).

[4] For additional references on executive compensation and franchising, see Prendergast (2002).

[5] See Allen and Lueck (2003) and Braido (2006) for further references on the tenancy literature.

risk tolerance. After addressing the bias caused by the endogenous matching of landlords and tenants, their results support risk sharing as an important determinant of sharecropping in early Renaissance Tuscany.

The data set used here allows one to precisely control for preference heterogeneity by comparing the different contracts of each agent. Contrary to Ackerberg and Botticini (2002) and consistent with the other papers on sharecropping, my findings do not support the existence of a negative correlation between risk and incentives.

# 3 Data

The data set is part of the village level studies (VLS) conducted by the International Crops Research Institute for Semi-Arid Tropics (ICRISAT), from 1975 to 1984, in eight different villages in India. The villages were selected to represent major agroclimatic zones. Initially, six villages were sampled in two different states: Aurapalle and Dokur (in the state of Andhra Pradesh); and Kanzara, Kinkheda, Shirapur, and Kalman (in the state of Maharashtra). Later, in 1980, the villages of Boriya Becharji and Rampura (in the state of Gujarat) were also included in the study.

For each village, ten households were randomly chosen in each of the following categories: large, medium, and small farmers, as well as landless workers. Households that emigrated from the villages were randomly replaced by another household in the same category. Resident investigators belonging to the same linguistic group as the villagers collected information on farming activities in each of the plots cultivated by the households. These investigators had rural backgrounds and their work was supervised by an economist from the ICRISAT.

Households typically cultivate multiple plots during each season of the year. Although the household is the primary sampling unit in the ICRISAT research, the schedule used here (the PS files) is organized at the plot level. Each observation refers to one of the plots cultivated by a household during a particular year and season. The panel is not balanced since farmers crop different plots over time (i.e., plots are not always observed through many periods). Moreover, there are plots that produce no output in some periods. These are likely to be plots under rotation or temporarily abandoned and, therefore, are not included in the study.

The sample comprises of 10,704 productive plots managed by 275 different households between 1975 and 1984. Table 1 describes the variables available for analysis. Notice that there are many products and byproducts being produced in each plot, but one type of culture is usually predominant (this culture is defined as being the main crop and identified by the first letter in the cropping-pattern code). Cereals, oil seeds, pulses, and fiber crops are the most common cultures. Moreover, the plots are mainly cultivated during the two seasons from June to February (i.e., the monsoon and post-monsoon seasons).

**[Table 1]**

It is important to stress how values were computed by the ICRISAT investigators. The actual value paid for seeds, fertilizers, pesticides, and manures and the rental value of rented bullocks and machinery (such as pumpsets and tractors) were recorded for each plot and season of the year. For home produced inputs and owned bullocks and machinery, the values were computed by multiplying the actual quantities used in each plot by village-specific prices and rents. Similarly, the database contains the actual value paid for hired labor, but the value of family labor is computed by multiplying the village wages for children, male adults, and female adults by the number of hours worked by each member in each plot. Finally, the value of the main product and byproducts were recorded at the prevailing village prices at the time of harvest.[6] Tables 2 presents the summary statistics for the quantitative variables.

**[Table 2]**

Table 3 compares the coefficient of variation of the per acre output value across lands managed under ownership, sharecropping, and fixed rent. It shows that, at least before controlling for the agents' characteristic, risk does not seem to be correlated with the land-contract format. The coefficient of variation is similar across lands under the three contract forms. Owner-operated lands present higher quality present higher quality, and their productivity mean and standard deviation are proportionally higher.

**[Table 3]**

## 3.1 Households under Multiple Contracts

A key characteristic of the database is the presence of households cropping multiple plots under different contracts. They are present in all the eight villages, but these households are mostly concentrated in the villages of Shirapur, Kalman, Kanzara, Boriya, and Rampura. The distribution of these plots across the years,

[6] For additional details, see Singh, Binswanger, and Jodha (1985).

seasons, and crops are similar to the full-sample pattern. There are households cropping lands under different contracts over time as well as households cropping lands under different contracts during the same period.

**Tenants with Different Contracts over Time.** There are 6,862 plot-level observations from households who have managed other plots under a contract with different incentive power over the periods. Among the 10,704 plots sampled, one has: 3,706 plots managed by households who only had lands under ownership over the periods (pure owner); 136 plots managed by households who only had lands under sharecropping over the periods (pure sharecropper); 4,313 plots managed by households who own and sharecrop different plots over the periods (mixed owner-sharecropper); 942 plots managed by households with plots under ownership and fixed rent over the periods (mixed owner-renter); 7 plots managed by households with plots under sharecropping and fixed rent over the periods (mixed sharecropper-renter); and 1,600 plots managed by households with lands under ownership, sharecropping, and fixed rent over the periods (mixed owner-renter-sharecropper).[7]

**Tenants with Different Contracts in a Same Period.** There are also 3,539 plot-level observations from households who have managed other plots under a contract with different incentive power in the same period (year and season). Among the 10,704 productive plots sampled, one has: 6,876 plots managed by households who only had lands under ownership during that particular period (pure owner); 252 plots managed by households who only had lands under sharecropping during that particular period (pure sharecropper); 37 plots managed by households who only had lands under fixed rent during that particular period (pure renter); 2,833 plots managed by households who own and sharecrop different plots during that particular period (mixed owner-sharecropper); 456 plots managed by households with plots under ownership and fixed rent during that particular period (mixed owner-renter); 5 plots managed by managed by households with plots under sharecropping and fixed rent during that particular period (mixed sharecropper-renter); and 245 plots managed by households with lands under ownership, sharecropping, and fixed rent during that particular period (mixed owner-renter-sharecropper).

[7] There is no plot managed by a farmer who only had lands under fixed rent over the years (pure renter).

# 4 Econometric Methodology

Changes in market prices and shocks in production are the two types of risk faced by the households. Prices are exogenous under the assumption that markets are competitive. However, final production could be affected by households' endogenous actions, which must be taken into account when measuring the exogenous risk.

As usual in the literature of agricultural economics, production is modeled here through a stochastic Cobb-Douglas function with constant return to scale. For each observation $i$ (a plot managed by some household in a certain year and season), one has:

$$y_i = a_i T_i^{(1-\alpha_k-\alpha_l)} k_i^{\alpha_k} l_i^{\alpha_l} \exp(u_i); \tag{1}$$

where $y_i$ represents the value of the final output; $T_i$ is the cropped area; $k_i$ and $l_i$ represent the value of nonlabor and labor input used; $a_i$ is a technological factor that accounts for household and land characteristics as well as specific effects associated with different villages, years, seasons, cropping pattern, and land contract; and $u_i$ is an error term accounting for exogenous shocks affecting prices and output. We assume that $E(\exp(u_i)) = 1$. This is without loss of generality since the factor productivity $a_i$ can vary across plots.

## 4.1 Identifying the Exogenous Risk

In this context, the input choices ($k_i$ and $l_i$) are endogenous and the error term $u_i$ captures all exogenous stochasticity in production. In a competitive environment without externality, the Pareto optimal input allocation should solve the following problem:

$$\max_{k_i, l_i} E\left(a_i T_i^{(1-\alpha_k-\alpha_l)} k_i^{\alpha_k} l_i^{\alpha_l} \exp(u_i) - k_i - l_i\right), \ \forall i. \tag{2}$$

Then, the optimal amount of nonlabor and labor inputs should given by:

$$k_i^* = T_i \left(\alpha_k^{(1-\alpha_l)} \alpha_l^{\alpha_l} a_i\right)^{\frac{1}{1-\alpha_k-\alpha_l}}; \tag{3}$$

$$l_i^* = T_i \left(\alpha_l^{(1-\alpha_k)} \alpha_k^{\alpha_k} a_i\right)^{\frac{1}{1-\alpha_k-\alpha_l}}. \tag{4}$$

Using this same data set, Braido (2008) tested and did not reject the hypothesis that input use is efficient across farms under different land contracts. Under this hypothesis, equation (1) could then be written as:

$$y_i = T_i \left[(\alpha_k)^{\alpha_k} (\alpha_l)^{\alpha_l} a_i\right]^{\frac{1}{1-\alpha_k-\alpha_l}} \exp(u_i). \tag{5}$$

Since the technological factor $a_i$ is poorly measured, one could also be interested in the following alternative representation for the production function. Notice from (4) that:

$$a_i = \frac{\left(\frac{l_i^*}{T_i}\right)^{(1-\alpha_k-\alpha_l)}}{\alpha_l^{(1-\alpha_k)}\alpha_k^{\alpha_k}}. \tag{6}$$

Therefore, equation (5) can also be written as:

$$y_i = \frac{1}{\alpha_l} l_i^* \exp\left(u_i\right). \tag{7}$$

The log-linear version of (7) is:

$$\ln\left(y_i\right) - \ln\left(l_i^*\right) = \ln\left(\frac{1}{\alpha_l}\right) + u_i. \tag{8}$$

Since detailed data on labor supply is available, one can identify $u_i$ from the residuals of an OLS regression of $\ln\left(y_i\right) - \ln\left(l_i^*\right)$ on a constant term.[8]

## 4.2 Testing the Incentive-Risk Relation

The error term $u_i$ accounts for exogenous shocks such as climatic changes, infestations, and blights. One can then use the heteroscedasticity test (see Glesjer, 1969) to assess whether the variance of $u_i$ differs across plots under different contracts. It is important to stress that the variance of the error term may depend on plots characteristics, such as the soil type and predominant crops. Therefore, the identification assumption in place here is that all characteristics affecting this plot-specific measure of risk are known by landlords and tenants at the time of setting the land contract.

Formally, consider the following class of models:

$$u_i^2 = g\left(d_i\beta + z_i\varphi + v_i\right), \tag{9}$$

where $g\left(\cdot\right)$ is a continuous and increasing function; $d_i$ is a vector with the relevant contract dummies; $z_i$ is a vector with household dummies to control for the

[8] One could also use the nonlabor input to measure $u_i$. However, given the large number of different nonlabor factors (see Table 1), this variable is more likely subject to measurement errors. The OLS estimation of equation $\ln(y_i) = c_0 + c_1 \ln(l_i^*) + u_i$ presents $\hat{c}_1 = 1.068$ (with clustered std. dev. equals 0.013). Analogously, the OLS estimation of equation $\ln(y_i) = c_0 + c_1 \ln(k_i^*) + u_i$ presents $\hat{c}_1 = 0.860$ (with clustered std. dev. equals 0.012).

household's risk aversion; and $\upsilon_i$ is an error term. The vector of parameters $\beta$ indicates how the different land contracts are correlated to the exogenous risk, given the characteristics of the household.[9]

# 5 Empirical Results

For the commonly used linear and quadratic specifications of the function $g(\cdot)$, the heteroscedasticity model becomes (respectively):

$$u_i^2 = d_i\beta + z_i\varphi + \upsilon_i, \tag{10}$$

and

$$\mid u_i \mid = d_i\beta + z_i\varphi + \upsilon_i. \tag{11}$$

In each case, the vector $\beta$ measures whether the variability of $u_i$ differs across lands under different contracts. Table 4 presents the results. The linear and quadratic specifications of the models are estimated three times: (i) without fixed effects; (ii) with household fixed effects; and (iii) with fixed effects per household and period (season of each year). Overall, risk is statistically equal across lands under fixed rent and sharecropping and slightly higher in owner-operated lands. This does not support the risk-sharing argument.

**[Table 4]**

# 6 Robustness

## 6.1 Subsample of Leased Plots

It is also interesting to investigate the validity of the risk-sharing prediction in a restricted subsample containing only leased plots. If owners and tenants accessed the same technology, these results should not differ from those based on the entire sample. In this case, restricting attention to the subsample of leased plots would only reduce efficiency of the estimates. However, if owners and tenants accessed different technologies, then the first-step regression and the risk measure could be modified when one restricts attention to the subsample of leased plots. Checking this possibility is the main motivation behind the exercise performed next.

[9] Note that linear regressions do not compute causal relationships, but only conditional correlations. Here, in particular, the exogenous term $u_i^2$ appears in the left-hand side of equation (9), while the endogenous contract dummies appear in the right-hand side of (9).

This subsample contains 1,796 leased plots, cropped by 134 different households. Around 88.7% of these plots are sharecropped and 11.3% of them are leased under fixed rent. Moreover, there are 479 plots being cropped by households who manage multiple plots under sharecropping and fixed rent during the years of data collection (allowing one to introduce household fixed effects). Finally, there are also 126 plots cropped by tenants that manage multiple plots under sharecropping and fixed rent during the same period (allowing one to introduce household-per-period fixed effects). The procedure based on the Glesjer's heteroscedasticity test is replicated for this subsample. The results, shown in Table 5, confirm that the exogenous error term is homoscedastic across plots under sharecropping and fixed rent.

**[Tables 5]**

### 6.2 Results Per Type of Crop

Allen and Lueck (1992) argue that the costs of dividing the output affect the design of land-leasing contracts. Since these costs vary across different cropping cultures, it is worth investigating whether the risk-sharing prediction is valid for some isolated cultures.

The Glesjer's heteroscedasticity tests are then performed for subsamples of plots cultivating each of the five most common cultures—namely, cereals; oilseeds; pulses; fiber crops; and vegetables and spices. One different production function is estimated for each of these cultures and the corresponding risk measures are constructed.

Thirty different testing procedures are implemented (relative to six regressions as in Table 5, for each of the five cultures). I do not report these results here. Among them, there are only three isolated regressions (for different cultures) in which either the ownership dummy or the fixed-rent dummy were statistically negative at the 5% level. Overall, the risk-sharing prediction is systematically rejected.

## 7 Conclusion

This paper tests a classic prediction of the moral hazard theory: the negative correlation between the contract's incentive power and its underlying risk, for each given agent. The empirical literature on contract theory relies on contracts from different agents to test this prediction. Ackerberg and Botticini (2002) point out for potential endogeneity bias inherent to this strategy. Here, the risk-sharing prediction is assessed by means of a database on tenancy contracts that contains households cropping multiple plots under different contracts. This feature allows us to compare the risk of different plots cropped by the same agent under different contracts. The findings do not support the existence of a negative risk-incentive relationship.

## References

[1] **Ackerberg, Daniel A. and Botticini, Maristella.** "Endogenous Matching and the Empirical Determinants of Contract Form." *Journal of Political Economy*, June 2002, 110 (3), 564-591.

[2] **Allen, Douglas and Lueck, Dean.** "Contract Choice in Modern Agriculture: Cash Rent versus Cropshare." *Journal of Law and Economics*, October 1992, 35 (2), 397-426.

[3] **Allen, Douglas and Lueck, Dean.** "Transaction Costs and the Design of Cropshare Contracts." *RAND Journal of Economics*, Spring 1993, 24 (1), 78-100.

[4] **Allen, Douglas W. and Lueck, Dean.** "The Role of Risk in Contract Choice." *Journal of Law, Economics, and Organization*, October 1999, 15 (3), 704-736.

[5] **Allen, Douglas W. and Lueck, Dean.** "The Nature of the Farm: Contracts, Risk and Organization in Agriculture." MIT Press, Cambridge, MA, 2003.

[6] **Aggarwal, Rajesh K. and Samwick, Andrew A.** "The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation." *Journal of Political Economy*, February 1999, 107 (1), 65-105.

[7] **Braido, Luis H.B.** "Insurance and Incentives in Sharecropping." In *Competitive Failures in Insurance Markets*, P.A. Chiappori and C. Gollier (eds.), MIT Press, Cambridge, MA, 2006.

[8] **Braido, Luis H.B.** "Evidence on the Incentive Properties of Share Contracts." *Journal of Law and Economics*, May 2008, 51 (2), 327-349.

[9] **Dee, Carol Callaway, Lulseged, Ayalew and Nowlin, Tanya S.** "Executive Compensation and Risk: The Case of Internet Firms." *Journal of Corporate Finance*, December 2005, 12 (1), 80-96.

[10] **Garen, John E.** "Executive Compensation and Principal-Agent Theory." *Journal of Political Economy*, December 1994, 102 (6), 1175-1199.

[11] **Glesjer, H.** "A New Test for Heteroscedasticity." *Journal of the American Statistical Association*, March 1969, 64 (325), 316-323.

[12] **Holmstrom, Bengt and Milgrom, Paul.** "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, March 1987, 55 (2), 303-328.

[13] **Jin, Li.** "CEO Compensation, Diversification, and Incentives." *Journal of Financial Economics*, October 2002, 66 (1), 29-63.

[14] **Lafontaine, Francine.** "Agency Theory and Franchising: Some Empirical Results." *RAND Journal of Economics*, Summer 1992, 23 (2), 263-283.

[15] **Lafontaine, Francine and Bhattacharyya, Sugato.** "The Role of Risk in Franchising." Journal of Corporate Finance, October 1995, 2 (1-2), 39-74.

[16] **Lambert, Richard and Larker, David.** "An Analysis of the Use of Accounting and Market Measures of Performance in Executive Compensation Contracts." *Journal of Accounting Research*, Supplement 1987, 25 (Studies on Stewardship Uses of Accounting Information), 85-125.

[17] **Prendergast, Canice.** "The Tenuous Trade-off between Risk and Incentives." *Journal of Political Economy*, October 2002, 110 (5), 1071-1102.

[18] **Rao, C.H. Hamunantha.** "Uncertainty, Entrepreneurship, and Sharecropping in India." *Journal of Political Economy*, May-June 1971, 79 (3), 578-595.

[19] **Singh, R.P., Binswanger, H.P. and Jodha, N.S.** "Manual of Instructions for Economic Investigators in ICRISAT's Village Level Studies." Technical Report, International Crops Research Institute for Semi-Arid Tropics, May 1985.

**Table 1. Data Description**

| Variable | Description |
|---|---|
| Output | Value of main output and byproducts, measured in Indian rupees |
| Ownership Dummy | 1 if plot is owned (83.2%); 0 otherwise |
| Fixed-Rent Dummy | 1 if plot is rented on a fix-rent basis (1.9%); 0 otherwise |
| Area Cropped | Actual area cropped, measured in acres |
| Nonlabor Input | Value of seeds, fertilizers, pesticides, organic and inorganic manures, plus the rental value of bullocks and machinery (in rupees) |
| Labor Input | Value of family and hired labor (in rupees) |
| Per Acre Land Value | Per acre value of the plot (in 100 rupees per acre) estimated by ICRISAT's investigators using information about potential sale value, topography, location, etc., obtained from a village specialist |
| Irrigation Dummy | 1 if the plot is irrigated (31.8%) |
| Soil-Type Dummies | 7.1% deep black; 34.3% medium black; 21.7% shallow black; 11.1% shallow red; 2.4% gravelly; 0.5% problem soil (saline, etc.); 9.8% sandy soil; 1.1% other soils; 12% undefined |
| Cropping Pattern | Qualitative variable (with 1,031 different codes) describing all products cropped in each plot |
| Main-Crop Dummies | Dummy variables constructed from the first letter of the cropping pattern code (which describes a general category for the dominant cropping product): 16.8% oilseeds; 53.2% cereals; 9.3% fiber crops; 0.4% garden crops; 14% pulses; 0.8% sugar cane; 4.2% vegetables and spices; 1.3% fodder crops |
| Village Dummies | 14.4% Aurepalle; 5.5% Dokur; 20.2% Shirapur; 15.7% Kalman; 14.6% Kanzara; 5.6% Kinkheda; 8.7% Boriya; 15.3% Rampura |
| Year Dummies | 1975 (10.9%); 1976 (11.1%); 1977 (10.3%); 1978 (9.7%); 1979 (9.5%); 1980 (9.2%); 1981 (10.6%); 1982 (9.9%); 1983 (9.5%); 1984 (9.3%) |
| Season Dummies | 35.8% planted from June to October; 58.5% from November to February; 5.5% from March to May; 0.2% perennial crops |
| Household | Village-specific numerical code that identifies the household |

*Note:* Data from the PS files of the Village Level Studies of the International Crops Research Institute for Semi-Arid Tropics (ICRISAT). The primary sampling unit is the household, but the observations refer to plots managed by each household during each season of the year.

**Table 2. Summary Statistics**

| Variable | Mean | Min. | Max. | St. Dev. | Sample Size |
|---|---|---|---|---|---|
| Per Acre Output | 754.1 | 0.68 | 24,964 | 1,106 | 10,704 |
| Per Acre Nonlabor Input | 318 | 0 | 16,478.8 | 507.2 | 10,704 |
| Per Acre Labor Input | 150 | 0.29 | 3,064 | 181.6 | 10,704 |
| Per Acre Land Value | 34 | 0 | 160 | 24.6 | 10,704 |

*Note:* Data from the ICRISAT's Village Level Studies.

**Table 3. Summary Statistics by Land Contract**

| | Per Acre Output | | | |
|---|---|---|---|---|
| | Coefficient of Variation (σ/μ) | Standard Deviation (σ) | Mean (μ) | Sample Size |
| **Ownership** | 1.45 | 1,166.5 | 804.5 | 8,908 |
| **Fixed Rent** | 1.23 | 722.4 | 587.0 | 203 |
| **Sharecropping** | 1.37 | 678.0 | 493.5 | 1,593 |

*Note:* Data from the ICRISAT's Village Level Studies.

**Table 4. Heteroscedasticity Test**

*Ordinary Least Square*

| | **No Fixed Effect** | | **Household Fixed Effects** | | **Household-Period Fixed Effects** | |
|---|---|---|---|---|---|---|
| | $(\hat{u}_{it})^2$ | $\lvert \hat{u}_{it} \rvert$ | $(\hat{u}_{it})^2$ | $\lvert \hat{u}_{t} \rvert$ | $(\hat{u}_{it})^2$ | $\lvert \hat{u}_{it} \rvert$ |
| **Ownership Dummy** | 0.050 | 0.022 | **0.080*** | **0.039*** | 0.079 | 0.038 |
| Robust Std. Err. | (0.039) | (0.019) | (0.043) | (0.022) | (0.051) | (0.027) |
| **Fixed-Rent Dummy** | 0.104 | 0.036 | -0.120 | -0.048 | -0.137 | -0.058 |
| Robust Std. Err. | (0.144) | (0.064) | (0.132) | (0.042) | (0.152) | (0.049) |
| **Constant** | **0.535***** | 0.547*** | Yes | Yes | Yes | Yes |
| *Sample Size* | *10,704* | *10,704* | *10,704* | *10,704* | *10,704* | *10,704* |

*Note:* The asterisks ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively. The cluster method is used to compute robust t-statistics and standard errors in order to account for the fact that households, rather than plots, are the primary sampling unit. Household fixed effects refer to 275 dummy variables generated from codes identifying the household (constant over different periods). Similarly, household-period fixed effects refer to 2,773 dummy variables generated through the iteration of codes identifying the household and the period (year and season).

**Table 5. Heteroscedasticity Test – Subsample of Leased Plots**

*Ordinary Least Square*

| | **No Fixed Effect** | | **Household Fixed Effects** | | **Household-Period Fixed Effects** | |
|---|---|---|---|---|---|---|
| | $(\hat{u}_{it})^2$ | $\lvert\hat{u}_{it}\rvert$ | $(\hat{u}_{it})^2$ | $\lvert\hat{u}_{t}\rvert$ | $(\hat{u}_{it})^2$ | $\lvert\hat{u}_{it}\rvert$ |
| **Fixed-Rent Dummy** | 0.089 | 0.023 | -0.083 | -0.001 | -0.167 | -0.007 |
| Robust Std. Err. | (0.135) | (0.060) | (0.203) | (0.068) | (0.265) | (0.103) |
| **Constant** | **0.534*****| **0.552***** | Yes | Yes | Yes | Yes |
| *Sample Size* | *1,796* | *1,796* | *1,796* | *1,796* | *1,796* | *1,796* |

*Note:* The asterisks ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively. The cluster method is used to compute robust t-statistics and standard errors in order to account for the fact that households, rather than plots, are the primary sampling unit. Household fixed effects refer to 134 dummy variables generated from codes identifying the household (constant over different periods). Similarly, household-period fixed effects refer to 663 dummy variables generated through the iteration of codes identifying the household and the period (year and season).