Working Paper Series

Theory of the Firm: Applied Mechanism Design

WP 96-02

Edward Simpson Prescott
Federal Reserve Bank of Richmond

Robert M. Townsend
Federal Reserve Bank of Richmond

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Working Paper 96-2

Theory of the Firm: Applied Mechanism Design*

Edward Simpson Prescott

and

Robert M. Townsend

Research Department
Federal Reserve Bank of Richmond

June 12, 1996

*Prescott is an economist in the Research Department at the Federal Reserve Bank of Richmond. Townsend is a professor in the economics department at the University of Chicago and is affiliated with the Federal Reserve Bank of Chicago. The authors acknowledge comments from participants at seminars at CREST, the Federal Reserve Bank of Richmond, LSE, and Yale. Both authors would like to thank ICRISAT for their help and support during visits to India. Prescott would like to thank the Jacob K. Javits Fellowship Program of the U.S. Department of Education and the Mellon Foundation for financial support. Townsend would like to thank the NSF and NIH for financial support. The views expressed here are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Chicago, the Federal Reserve Bank of Richmond, or the Federal Reserve System.

Abstract

This paper studies the question: Why are there Firms? Motivated by observations of a variety of economies, several distinct concepts of what it means to be a firm are identified and then analyzed with mechanism design models. In the first class of models, a group of individuals is a firm if they collude and share information. This model is analyzed and compared with the non-firm alternative. Conditions are provided in which firms are preferred to no firms and vice versa. Next, we show how an economy with multiple distinct groups of colluding individuals can be decentralized.

In the next class of models, collusion is prohibited, but the information structure of the economy depends on whether individuals work together. Activities performed jointly by the individuals are considered to be done within a firm. In the model the degree of economic activity organized by firms is endogenous. Numerical examples are provided in which none, some, and all work is done within a firm. The last class of models studies the long-term nature of firms versus the short-term nature of non-firm arrangements, like spot markets. Multi-stage models are developed in which individuals may be switched between projects. People who work the same project over time are considered to work for a firm. Conditions for the optimality of long-term and short-term arrangements are provided.

1 Introduction

Consider observations from five different economies:

1. In and around the southern Indian village of Aurepalle, located in the Mahbubgnar district of Andhra Pradesh, there is a relatively common form of economic organization. As described in Mueller and Townsend (1994), a group of farmers, often unrelated to one another, agree to jointly farm the land of a relatively large landholder. Each farmer promises to contribute labor and pairs of bullocks and to share profits or losses with one another in proportion to the number of pairs provided. They promise to work the rented plots together throughout a cropping season and to share labor and bullock burdens without other compensation. Landowners not only contract with cropping groups but more often contract with individual tenants, sometimes dividing their landholdings to do so. Indeed, not all land is brought into tenancy; both landowners and cropping group members often farm some of their own land separately. The village even contains alternatives to tenancy. Other farmers hire distinct laborers or bullock operators for specific operations one at at time, as in an active spot market. For a description of the market for short-term bullock services see Mueller and Prescott (1996).

2. Indian cropping groups may seem more familiar when viewed in a different context. Professional partnerships in the U.S. are a similar type of firm. For example, a group of senior lawyers form a firm or group which divides profits or losses ex post and competes in the market for lawyer services. Physician and legal partnerships have been the subject of modeling by Gaynor and Gertler (1996) and Lang and Gordon (1996), respectively.

3. The "family" might also be thought of as an alternative, more "communal" group, as in Chiappori (1992), but caution should be exhibited when interpreting the family as communal. In villages in Burkina Faso a given family may divide its plots of land, even among its own members. Udry (1994) describes families where male members operate plots distinct from female members and otherwise identical plots are farmed differently by each gender. There may be rules for how much of the output of given plots is kept by the individual members and how much is shared within the larger family group.

4. In villages of Thailand some families combine together along larger kinship lines, sharing land, consumption, and labor. These extended networks interact with the larger village economy, in a variety of ways, including in exchange arrangements, in spot markets, and in various forms of credit and insurance. The boundary defining what is done in the family and what is done outside it, that is, what is done in firms and what is done in markets, are not yet clear.

5. Finally, in the altiplano of Bolivia the Aymara Indians designate certain plots as communal. These are allocated by tribal leaders to participating families from year to year, sometimes by lot. Communal plots are to be worked at the same time, on specified days, by all participating families. Other plots are held year after year along family lines and worked with more autonomy.

These observations describe a variety of institutions which are like firms. The institutions are characterized by groups of people working together under an agreed upon set of rules. Further, these firms have been chosen over and may coexist with various non-firm alternatives including individual ownership, individual tenancy, and spot markets. This paper provides theoretical models which attempt to address these observations. The question we ask is the fundamental one posed in Coase (1937): Why are there firms?

Firms are studied by identifying several distinct concepts of what it means to be a group or firm, and writing down explicit models based on these concepts. In the first class of prototypes, studied in section 2.2, a group of individuals, or households, is said to be a firm if they share information and collude among themselves in their relationship with a landowner or the rest of the village. By collusion we mean that the group solves a Pareto problem among themselves, and economy-wide allocations must respect their ability to do this. The idea is that members of firms jointly make decisions and will do so in their best interests.

In section 2.3, the collusive arrangement is compared with a non-firm alternative in which individual tenancy or ownership is imposed a priori. Section 2.1 analyzes the individual tenancy, non-firm, regime which is characterized by a lack of within-group information-sharing and a lack of collusion. The dominance of one arrangement over another is a function of the distribution of wealth and other parameter values, which is consistent with the observed multiplicity of organizational arrangements. The results extend the relative performance literature, started by

Holmström and Milgrom (1990), Ramakrishnan and Thakor (1991), and Itoh (1993).

Although this paper often refers to groups as collusive, communal entities, we only mean that a group decides on its own rules for an internal resource allocation given the group's interaction with the rest of the economy. Rules need not be of the command and control form, as used in a direct mechanism. In particular, section 2.4 uses the Second Welfare Theorem to show that under certain conditions within-group allocations can be decentralized by a price system. The main result of the section, however, is to show that a multiple-firm economy can be decentralized by allowing firms (the groups) to buy and sell commodities in a competitive market. Our analysis extends the work of E.C. Prescott and Townsend (1984a) and (1984b) on decentralization and private information to an environment where there may be collusion within groups.

We study as well a second class of prototypes motivated by a different definition of a firm. In section 3, we prohibit collusion among farmers and instead focus on conditions under which farmers would jointly farm plots of land. Jointly farmed plots are advantageous because, without collusion, farmers' efforts on these plots can be revealed publicly. In contrast, efforts on individually-farmed plots are presumed to be unobserved by other farmers. Costs to jointly farming plots are allowed as well so the optimal land assignment trades off these costs with the information benefit just mentioned.

In this second prototype communal plots can coexist with individual ownership and with single-agent tenancy. Single-agent tenancy still allows cross-household resource allocation and cross-household insurance but under the more stringent constraints resulting from private information of efforts on individually operated plots. As with the first class of prototypes the appropriate land assignment depends on parameter values.

The last class of prototypes focus on a third aspect of firms. Section 4 studies season-long assignment of labor to plots and compares this with reallocating labor across plots of land by task-specific assignment, as would happen in a spot market. As might have been anticipated, season-long contracts can be good for incentives. However, in some versions of the models, task-specific assignments and labor reallocations are also good for incentives and allow for a Pareto superior allocation of resources. Again, the appropriate choice between season-long contracts and spot markets depends on the parameter values and in particular on the nature of the production technology.

As is evident, the theoretical analysis breaks our study into several distinct categories. The categorization is done for clarity. In practice multiple features of firms may be appropriate for a given economy. Certainly, prototypes can be combined. The second prototype, assigning farmers to plots of land, could be extended to allow collusion, and the third prototype could be extended to allow long-term contracts within groups.

2 Firms as Cooperatives

For the first class of models imagine that there is a set of n farmers who can work plots of land. Efforts on plots of land determine the probability distribution of outputs. The farmers can potentially come together to constitute a cooperative or firm relative to an outsider who can be either a landowner or the rest of the economy. Two regimes are considered and compared. In the first regime the n farmers have full information about each other's efforts, outputs and consumption and they can collude against the outsider by agreeing to an internal effort and consumption allocation which is Pareto optimal for the group, conditioned on the group's agreement with the outsider. We call this setup the firm or group regime.

In the second regime two changes are made. First, collusion between the farmers is prohibited. The outsider or principal can prevent any reallocation of effort or consumption. He determines and knows all transfers. Second, farmers no longer observe each other's efforts. Instead, they only observe what the principal does, namely, the other farmer's outputs. Models with these assumptions are called relative performance models.

This section explicitly compares, under varying technological assumptions, the group regime with the relative performance regime. The effect on allocations of each organizational form is described and compared. Examples are provided which demonstrate how the parameters describing an economy alter the organizational choice. Implicit in the comparison is the assumption that if farmers do observe each other's efforts then they have the ability to collude.[1]

Drawing on the agrarian economies described in the introduction, imagine that there is one plot of land available to be worked for each farmer i of the potential group. For simplicity assume

[1]These assumptions need not be bundled together. For example, Hammond (1987) and Haubrich (1988) study models where agents side-trade with each other despite private information. In our context, collusion would add constraints to the relative performance program making the organizational choice trivial. Consequently, we do not pursue that class of models in this section. However, there are organizational design models in which the assumptions can be unbundled and a bit more will be said about this later.

that there are only two farmers and two plots of land. Each farmer i has preferences over own consumption, c_i, and total own effort, e_i. Total own effort is the sum of farmer i's efforts across all plots $j = 1, 2$. Let e_{ij} denote farmer i's effort on plot j, so $e_i = \sum_j e_{ij}$. It is also useful to define $\mathbf{e_{i\bullet}} = (e_{i1}, e_{i2})$ as the *vector* across the two plots of farmer i's efforts. Utility is defined by $U(c_i, e_i)$.

Output on plot j, q_j, is a function of the total effort put on the plot, $a_j = \sum_i e_{ij}$, and a random shock. The plots of land could differ in size or productivity. The probability of outputs q_1 and q_2, given total plot efforts a_1 and a_2, is described by the density function $p(q_1, q_2|a_1, a_2)$. For expositional ease the same function will be often written $p(q_1, q_2|\mathbf{e_{1\bullet}} + \mathbf{e_{2\bullet}})$, because $\mathbf{e_{1\bullet}} + \mathbf{e_{2\bullet}} = (a_1, a_2)$. Finally, correlation in plot returns, from common shocks like rain or temperature, is easily incorporated into this production function.

Two different specifications of the technology regarding the plots are considered. In the first specification, farmers can work each other's plots, whether done in the group regime by colluding against the outsider, or done in the relative performance regime in assignments recommended by the outsider. In the second specification, farmers must only work their assigned plots. This technology is modeled by restricting efforts to satisfy $e_{12} = e_{21} = 0$.

It is useful now to connect the varying technological specifications to the example economies described earlier. Consider first the Indian cropping groups. The most appealing technological specification is that group members can work, as they choose, any or all of the group's plots. The second technological specification is more appealing when a landowner rents his land to different tenants (or to multiple groups), as does happen. Under this second specification, tenants could collude in efforts, thus constituting a group. Under either specification the group shares consumption risk internally.

Alternatively, in the relative performance regime, the landowner would recommend labor efforts on each of the various plots. In the first technology specification a given tenant could be asked to work any of the landowner's plots at any time, like the permanent servants of a landowner do. Somehow the owner controls each servant's consumption and precludes collusion among the servants. In the second technology specification, tenants are assigned to one plot all of the time. Here again, collusion is not presumed possible if the plots are far apart.

Also, consider the Aymara families assigned to work communal plots. Each plot might be farmed individually by a given family as in the second technology specification. One local leader claimed this was the case. Or, instead, each family could share labor with the other families, as in the first technology specification. In either specification participating families would share consumption risk and hence act as a group relative to its obligation to some outsider. The alternative to groups is that after land assignments families neither coordinate effort nor share consumption risk, and for us would not constitute a firm.

Likewise, individual members of a given family in Burkina Faso or given families of an extended kinship group in Thailand may work their own plots, as in the second technology specification. However, they might collude in paying off debts or rents to nonfamily or nonkinship group members. In this sense they act as a group. If they share labor as well, then the first technological specification is appropriate. The alternative to group-like arrangements is individual families operating in isolation, at least in the sense of not sharing information nor colluding.

Farmers in these economies are imagined to be dealing with an outsider, but this outsider has not yet been described. Two concepts of the outsider are useful. The first is that the outsider is a principal. In an Indian village the principal would be a wealthy landowner who does not supply labor, either because he is an absentee landowner, or because other activities are a more valuable use of his time. A landowner's utility is a function of plot outputs minus tenants' consumptions, namely $W(q_1 + q_2 - c_1 - c_2)$. The alternative concept of the outsider is that he is simply the rest of the economy, so W is linear and $q_1 + q_2 - c_1 - c_2$ is the surplus generated by a group. In a closed economy with multiple groups the summation of surpluses across groups must be zero.

The modeling strategy is to solve a class of planning problem for each regime and each technology specification. The solutions to each planning problem determines the set of Pareto optimal allocations for that case. A particular Pareto optimum would be determined by weights associated with the wealth and status of the individual farmers and of the outsider. The choice object in the programs will be a potentially random device which assigns consumptions $c_i \in C_i$ to both farmers, outputs $q_i \in Q_i$ on both plots, and effort of farmer i on plot j, $e_{ij} \in E_{ij}$, where upper case denote the set of feasible points of the relevant set. Recall that in the second technological specification, when farmers must only work their own plot, $E_{ij} = \{0\}$, $i \neq j$.

For expositional clarity define $\mathbf{c} = (c_1, c_2)$ and $\mathbf{q} = (q_1, q_2)$ as the vectors of consumptions and outputs, respectively. As mentioned earlier $\mathbf{e_{i\bullet}}$ is the vector of farmer i's plot-specific efforts. The choice object can now be written as $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$, the probability with which a particular consumption, output, and labor allocation is chosen.

Each program maximizes the weighted sum of the expected utilities of the two farmers subject to the following constraints: a participation constraint for the outsider; incentive constraints on individual efforts if it is a relative performance model, or a collusion constraint if it is a group model; nature constraints which assure that the endogenous probability of outputs is consistent with the underlying technology specification $p(q_1, q_2|\mathbf{e_{1\bullet}} + \mathbf{e_{2\bullet}})$; and a set of constraints which ensure that the choice object π is a probability measure. The advantage of the lottery approach is that when combined with assumptions that consumptions, efforts, and outputs must be chosen from a finite set, the constrained maximization problem can be written as a linear program. If the dimensions of the sets are small enough, solutions to parameterized economies can be computed.

2.1 Relative Performance Regime

It is useful to begin the discussion with the relative performance models. In these models one tenant cannot observe the other tenant's effort nor can the two tenants collude. Program 1 describes the problem under the technological specification that tenants can work either plot.

Program 1

$$\max_{\pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})} \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})[\lambda_1 U(c_1,e_1) + \lambda_2 U(c_2,e_2)]$$

$$\text{s.t.} \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) W(q_1 + q_2 - c_1 - c_2) \geq \bar{W}, \tag{1}$$

$$\begin{aligned}&\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) U(c_1,e_1)\\ &\quad\geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) \frac{p(\mathbf{q}|\mathbf{\hat{e}_{1\bullet}} + \mathbf{e_{2\bullet}})}{p(\mathbf{q}|\mathbf{e_{1\bullet}} + \mathbf{e_{2\bullet}})} U(c_1,\hat{e}_1), \quad \forall \mathbf{e_{1\bullet}}, \mathbf{\hat{e}_{1\bullet}},\end{aligned} \tag{2}$$

$$\begin{aligned}&\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) U(c_2,e_2)\\ &\quad\geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) \frac{p(\mathbf{q}|\mathbf{e_{1\bullet}} + \mathbf{\hat{e}_{2\bullet}})}{p(\mathbf{q}|\mathbf{e_{1\bullet}} + \mathbf{e_{2\bullet}})} U(c_2,\hat{e}_2), \quad \forall \mathbf{e_{2\bullet}}, \mathbf{\hat{e}_{2\bullet}},\end{aligned} \tag{3}$$

$$\forall \mathbf{\bar{e}_{1\bullet}},\ \mathbf{\bar{e}_{2\bullet}},\ \mathbf{\bar{q}},\ \sum_{\mathbf{c}} \pi(\mathbf{c},\mathbf{\bar{q}},\mathbf{\bar{e}_{1\bullet}},\mathbf{\bar{e}_{2\bullet}}) = p(\mathbf{\bar{q}}|\mathbf{\bar{e}_{1\bullet}} + \mathbf{\bar{e}_{2\bullet}}) \sum_{\mathbf{c},\mathbf{q}} \pi(\mathbf{c},\mathbf{q},\mathbf{\bar{e}_{1\bullet}},\mathbf{\bar{e}_{2\bullet}}) \tag{4}$$

$$\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) = 1, \text{ and } \forall \mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}},\ \ \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) \geq 0. \tag{5}$$

Constraint (1) is the participation constraint for the principal. By varying $\bar{W}$ and the λ weights the Pareto frontier can be calculated. Constraints (2) are the incentive constraints for tenant one. They imply that for every effort $\mathbf{e_{1\bullet}}$ assigned with positive probability, obedience is weakly preferred to deviations $\hat{\mathbf{e}}_{\mathbf{1\bullet}}$. As the summand indicates tenant one makes his decision independent of the other agent since he does not observe the other tenant's effort or recommendation. Similar to (2) are equations (3), tenant two's incentive constraints. Equations (4) guarantee that allocations respect the probability distributions generated by the technology. Finally, line (5) ensures that the choice variable is a probability measure.

Only one slight modification to Program 1 is required to set up the constrained maximization problem which incorporates the restriction that tenants must work their own plots: restrict the set of feasible allocations to those where $e_{12} = e_{21} = 0$. Everything else, the objective function and the constraints, are unchanged. The program for this technological specification will be referred to as Program 2, but because of the similarities with Program 1 it is not shown.

The advantage of the relative performance regime is that the outsider retains full control over consumption and uses this to induce efforts. Indeed, we can demonstrate this while taking up the issue of whether each agent would be assigned to work his own plot, or whether plots would farmed jointly. When outputs across the two plots have a common component or are otherwise correlated, then the outsider can infer from output on one plot whether the other farmer was shirking. This is particularly evident in the special case where the two agents have identical utility functions, identical plots, and equal Pareto weights. If returns on the plots are perfectly correlated for given efforts, then a simple contract implements the full-information solution: Assign both agents the same level of effort and make them work only their own plot. (Recall that in the first technological specification, agents can work either or both plots.) If plot outputs differ, then give both agents zero consumption; if outputs are identical then regardless of the level of plot outputs, reward each agent with a constant amount. If zero consumption provides a low enough level of utility, this arrangement, the full-information allocation, is incentive compatible. For small changes in the technology this result should be approximately true.

It should be clear why the principal does not want farmers to work more than one plot. If a farmer worked both plots then he could deviate from recommended efforts on both plots simultaneously and under the assumptions of this relative performance model neither the other tenant nor the principal would observe his shirking. Neither could the joint deviation be revealed by comparing outputs across plots since outputs would still be the same across plots. A joint deviation would change the probability distribution of outputs without altering the comovement of outputs across the plots.

Would it be desirable to assign agents to different plots if returns are uncorrelated? The answer: it depends. No longer can efforts on one plot be inferred from outputs on the other plot but if more than one agent is working a plot than there can be a detrimental effect on incentives, making individual assignments optimal. The following example demonstrates this effect for a particular class of equilibria. The example is worth noting because a variant of it will be used to address the optimality of long-term arrangements in section 4.

Consider an economy with two identical farmers and two identical plots, and Pareto weights $\lambda_1 = \lambda_2 = .5$. Farmers' sets of consumption, output, and efforts are equal and the effort grids are composed of equally spaced points. Plot returns are uncorrelated and the production technology on each plot satisfies the monotone likelihood ration property (MLRP) and the convexity of the distribution function (CDFC) conditions. When these conditions hold, the incentive constraints can be replaced with first-order conditions in the standard principal-agent problem. See, for example, Hart and Holmström (1985). In our model with grids of the underlying variables the equivalent of the first-order approach is to only check incentive constraints for downward deviations adjacent to the optimum.

In this environment, we restrict ourselves for the moment to analyzing *symmetric* equilibrium, that is, allocations where both plots are worked equal amounts and both farmers work equal amounts.

Theorem 1 *Under the technological specification where farmers may work either plot, if plot returns are uncorrelated and satisfy MLRP and CDFC then the optimal symmetric equilibrium is to assign each agents to a separate plot.*

Proof: First consider the assignment where each farmer only works his own plot. As Holmström (1982) has shown, if plot returns are uncorrelated then each farmers' compensation depends only on output on his own plot. Consequently, each tenant's relationship with the principal can be treated like a single-agent problem, so by the assumptions on the technology only downward adjacent constraints need to be checked. Let e_{11}^h index the hth element in the E_{11} grid. The incentive constraints for farmer one are then

$$\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})U_1(c_1,e_{11}^h)$$
$$\geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})\frac{p(q_1|e_{11}^{h-1})}{p(q_1|e_{11}^{h})}U_1(c_1,e_{11}^{h-1}). \tag{6}$$

Now consider an assignment where total plot effort is unchanged but each farmer works both plots. The assumptions on the production technology are no longer sufficient for the first-order approach to be valid, see for example Itoh (1991). However, the downward adjacent incentive constraints on each plot are a subset of the total constraints, so any allocation must satisfy them as well as satisfying the others. Letting h_j index elements of the E_{ij} grid, these constraints for tenant one are written

$$\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})U_1(c_1,e_{11}^{h_1}+e_{12})$$
$$\geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})\frac{p(q_1|e_{11}^{h_1-1}+e_{21})}{p(q_1|e_{11}^{h_1}+e_{21})}U_1(c_1,e_{11}^{h_1-1}+e_{12}), \tag{7}$$
$$\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})U_1(c_1,e_{11}+e_{12}^{h_2})$$
$$\geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})\frac{p(q_1|e_{12}^{h_2-1}+e_{22})}{p(q_1|e_{12}^{h_2}+e_{22})}U_1(c_1,e_{11}+e_{12}^{h_1-1})). \tag{8}$$

Constraints (6) and (7) are equivalent, but allocations satisfying (6) need not satisfy (8) let alone the incentive constraints not shown. Finally, it is feasible for the planner to randomly assign farmers to work separately or together. In this case a randomized assignment would simply entail an ex ante lottery over whether constraints (6) or constraints (7) and (8) would hold. This can not help since by the previous argument being constraint (6) with certainty would be less binding. **Q.E.D.**

Intuitively, if a farmer is assigned to both plots, his efforts affect the probability distribution of both plot outputs. At the margin his decrease in disutility from effort is the same regardless of which plot he works less. Consequently, equally strong incentives are required on both plots. In contrast if a farmer works only one plot, the incentives can be concentrated solely on the returns of a single plot.

While the above example provides intuition into the relative performance model, it uses a restrictive definition of equilibrium. We do not know if it is optimal in less restrictive classes of equilibria. As the next example demonstrates, assigning agents to separate plots is not optimal in general. It suggests that conditions such as concavity in technology and minimal inference benefits from increasing total effort on a plot are important in determining whether farmers are assigned to the same plot or not.

Consider a simple economy where on each plot there is only a low or high output. Let the probability of the high output be linear in the effort and let returns across plots be uncorrelated, that is, $p(q_j|e_{1j}+e_{2j}) = e_{1j}+e_{2j}$. Further assume that up to one unit of effort may be put on each plot so the probability of the high output is unity (no uncertainty) if one unit of effort supplied to that plot.

Rather than calculating the solution to an example we demonstrate that for certain levels of total effort, working the plots together is preferable to working them symmetrically. The implication is that there will be examples which have the characteristics of the proposition presented below.

Proposition 1 *Under the described technology there exists levels of total effort at which it is optimal for both agents to work together.*

Proof: Consider allocations where both agents work 0.5 units of effort. A symmetric allocation across plots requires tenant one to work 0.5 units of effort on the first plot and the other tenant to work 0.5 units on the other plot. This allocation requires dependence of consumption on output to induce each agent to work the desired amount. Consequently, with risk-averse preferences there is variability in consumption and the allocation deviates from the full-information solution.

Now consider an alternative allocation in which, again, both agents work 0.5 units of effort. Assign both agents to work the *same* plot. Let the compensation schedule be both agents receive consumption c if output on plot one is high and both receive zero if it is low. This contract may not look like a full-insurance contract because consumption depends on output, but the technology is such that the high output is produced with probability one *if* both agents work their recommended efforts. If they both do, then in equilibrium there is no variability in tenants consumption, just as in a full-information solution.

In this allocation a deviation by either agent, is detectable if output is low. It is punished by providing the lowest level of consumption to both agents. The principal must provide zero consumption to both agents since he does not know which agent deviated, but if the disutility of zero consumption in the low output state is low enough the contract is incentive compatible. With the caveat that the zero consumption is a strong enough punishment, the full-information characteristics of the allocation implies that working together is preferred to the working apart. **Q.E.D.**

In the symmetric example inference was improved by separating the tenants so that the principal could clearly determine who was responsible for which actions. The lesson of the other example is that incentives can also be reduced through the effect of aggregate plot efforts on inference. Itoh (1991), in a different environment, also observes that total effort can by asymmetrically applied to the two plots, though he focuses on the potential non-optimality of symmetric allocations rather than any inference benefits.

2.2 Group Regime

The paper now considers the case where farmers observe each other's efforts and are in a group which colludes in its dealings with the outsider. Consider first the specification where labor effort can be shared across the plots. The idea is to start with Program 1 above and then imagine what the two farmers would do if they could jointly specify aggregate and individual efforts on the two plots, could specify internal consumption as a function of outputs across the two plots, have perfect information about efforts as well as output, and have the ability to perfectly and costlessly enforce all internal agreements.

These assumptions change the information structure in the group model from that of the relative performance model. Here in particular the group members have perfect information about their labor efforts. The question naturally arises as to whether the outsider can take advantage of this using a direct revelation mechanism which implements the full-information optimum. If members were not allowed to collude then such an allocation could be implemented as in Harris and Townsend (1981). But as Itoh (1993) has argued in a similar context, collusion effectively rules this out. The two farmers would simply decide on the allocation they would like to implement at the beginning, after the contract with the principal has been signed. They would then send messages and adopt strategies to steer the revelation mechanism to this desired outcome. Consequently all the (unnecessary) notation which would accompany such direct revelation schemes is suppressed.

One can ask whether collusion among the agents could take place at some other point along the time line of the contract. For example, could the two farmers collude ex ante at the time the contract is signed by committing with the principal to carry out full-information allocations? As much as all parties to the contract might like this arrangement, the two farmers would re-solve given the full-information contract for a conditionally Pareto allocation in the group, much the way a single agent takes actions conditioned on the contract with the principal. This type of collusion cannot be prevented ex ante.

Ex post collusion after outputs are realized is feasible, but it accomplishes nothing new. Given that the two farmers have agreed to a conditionally Pareto optimal allocation and end up with specified consumption bundles as a function of realized outputs and the payoff to the principal, there is no possible gain for both farmers together. Changes which benefit one farmer must by construction hurt the other. Accordingly, we suppose the agents have commitment devices which preclude this kind of ex post bargaining.

In summary, the two farmers do all their bargaining ex ante, before efforts are taken and before outputs are realized, but after making an agreement with the outsider. It is assumed that the outcome of this bargain is a conditional Pareto optimal group allocation, one which maximizes a weighted sum of utilities. (The weights could come from different land holdings, time endowments, or other factors reflecting a priori diversity across these two farmers.)

To find a conditional Pareto optimal group allocation, one proceeds by considering the sub-problem of maximizing a sum of weighted utilities of the two farmers subject to a resource constraint that individual efforts must sum to the aggregate efforts which the group members agreed to do, and a resource constraint that internal consumptions add to the sum of consumption implicit in the agreement with the outsider to give or receive transfers as a function of observed output.

In particular, consider the problem faced by the agents after vectors of efforts $\mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}$ have been recommended by the outsider under the Program 1 allocation, $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$. The group, having decided in some way on internal weights, sharing consumption risk, and coordinating efforts, cares only about total resources it has available to distribute among its members. Let $c_g = c_1 + c_2$ denote total group consumption, and let $e_g = e_1 + e_2 = a_1 + a_2$ denote total group effort. Given $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$, the group faces from their perspective an implicit production function, $p(c_g|\mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$. It is easily calculated from the initial allocation, $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$:

$$p(c_g|\mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) = \sum_{\{c_1, c_2 | c_1 + c_2 = c_g\}} \pi(c_1, c_2|\mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) p(q|\mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) \tag{9}$$

Faced with this implicit production function, the group decides on $\tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet}$. Equivalently, they decide on total effort, $\tilde{e}_g$, effort across plots, $\tilde{a}_1$ and $\tilde{a}_2$, and effort of each group member, $\tilde{e}_1$ and $\tilde{e}_2$. Further, upon receiving c_g units of consumption the group divides it between its members by choosing, $\tilde{\mathbf{c}} = (\tilde{c}_1, \tilde{c}_2)$. If we let $\mu = (\mu_1, \mu_2)$ denote the planner's weights *within* the group, group members choose these efforts and consumptions to maximize the weighted sum of their utilities, subject to the group transfer rule (9) provided by the outsider. Varying μ is a way to trace the set of *within* group Pareto optimal allocations. The subproblem faced by the group is written:

Program S

$$\max_{\tilde{\pi}()} \sum_{\tilde{\mathbf{c}}, c_g, \tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet}} \tilde{\pi}(\tilde{\mathbf{c}}, c_g, \tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet})[\mu_1 U(\tilde{c}_1, \tilde{e}_1) + \mu_2 U(\tilde{c}_2, \tilde{e}_2)]$$

$$\text{s.t. } \forall \bar{c}_g, \;\; \bar{\tilde{\mathbf{e}}}_{\mathbf{1}\bullet}, \bar{\tilde{\mathbf{e}}}_{\mathbf{2}\bullet}, \;\; \sum_{\tilde{\mathbf{c}}} \pi(\tilde{\mathbf{c}}, \bar{c}_g, \bar{\tilde{\mathbf{e}}}_{\mathbf{1}\bullet}, \bar{\tilde{\mathbf{e}}}_{\mathbf{2}\bullet}) = p(\bar{c}_g | \bar{\tilde{\mathbf{e}}}_{\mathbf{1}\bullet}, \bar{\tilde{\mathbf{e}}}_{\mathbf{2}\bullet}) \sum_{\tilde{\mathbf{c}}, c_g} \pi(\tilde{\mathbf{c}}, c_g, \bar{\tilde{\mathbf{e}}}_{\mathbf{1}\bullet}, \bar{\tilde{\mathbf{e}}}_{\mathbf{2}\bullet}),$$

$$\sum_{\tilde{\mathbf{c}}, c_g, \tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet}} \pi(\tilde{\mathbf{c}}, c_g, \tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet}) = 1, \text{ and } \forall \tilde{\mathbf{c}}, c_g, \tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet} \;\; \pi(\tilde{\mathbf{c}}, c_g, \tilde{\mathbf{e}}_{\mathbf{1}\bullet}, \tilde{\mathbf{e}}_{\mathbf{2}\bullet}) \geq 0.$$

Apart from the potentially random choice of efforts, this program simplifies dramatically for the case of utility functions separable in consumption and effort. Let $U(c_i) - V(e_i)$ be the separable utility function, where U and V are strictly concave. Recall that tenants' efforts are perfect substitutes in production on either plot. If, for the moment, consumption and efforts are assumed to be chosen from a continuum, then the allocation of consumption $\tilde{c}_i$ and individual efforts $\tilde{e}_i$, $i = 1, 2$, in the group should satisfy first-order conditions:

$$\mu_1 U'(\tilde{c}_1) = \mu_2 U'(\tilde{c}_2), \tag{10}$$

$$\mu_1 V'(\tilde{e}_1) = \mu_2 V'(\tilde{e}_2). \tag{11}$$

Equations (10) and (11) can be solved numerically, or occasionally analytically, for internal distribution rules:

$$\tilde{c}_1 = \tilde{c}_1(c_g, \mu), \;\; \tilde{c}_2 = \tilde{c}_2(c_g, \mu), \tag{12}$$

$$\tilde{e}_1 = \tilde{e}_1(e_g, \mu), \;\; \tilde{e}_2 = \tilde{e}_2(e_g, \mu), \tag{13}$$

where $e_g = \sum_{ij} e_{ij} = a_1 + a_2$ is total group effort. If final consumptions and efforts are restricted to grids, as in the earlier programs, the internal distribution rules are computable even though the first-order conditions (10) and (11) are not be valid. For example, consumption sharing rules are computed by solving over the grid $P = C_1 \times C_2$ such that $c_1 + c_2 \leq c_g$ the problem $\max_{\pi()} \sum_{c_1,c_2} \pi(c_1, c_2) \sum_i \mu_i U_i(c_i)$ subject to π being a probability measure. The solution to this program, π^*, is only a function of c_g and μ, just like in equation (12). Individual efforts e_1 and e_2 are similarly determined. Consequently, for notational simplicity we use equations (12) and (13) in the following analysis.

Equations (10) and (11), or their discrete analogues, are important because they imply that only aggregate group consumption c_g and aggregate group effort e_g matter in the determination of the internal distribution of consumption and effort, given the μ weights. More will be said later about the relationship between individual allocations and group allocations.

Invocation of the revelation principle here would argue that the group's potentially random choice of efforts and their redistribution of consumption, can be anticipated and loaded into an initial random assignment function $\pi(\mathbf{c}, \mathbf{q}, a_1, a_2, \mu)$. Then, conditional on an assignment, recommended plot-specific efforts (a_1, a_2) should maximize the μ weighted sum of expected utilities relative to any other such choice. Similarly, the group's sharing rule for farmer-specific consumptions and efforts can be anticipated, by equations (10) and (11) for example. Consequently, aggregates c_g and $e_g = \sum_{ij} e_{ij}$ along with the *within-group* Pareto weight μ can be assigned so that individual consumptions and efforts are replaced with (12) and (13). Letting λ be the vector of planner weights the program is as follows:

Program 3

$$\max_{\pi()} \sum_{c_g,\mathbf{q},a_1,a_2,\mu} \pi(c_g, \mathbf{q}, a_1, a_2, \mu) \sum_i \lambda_i [U(c_i(c_g, \mu)) - V(e_i(e_g, \mu))]$$

$$\text{s.t.} \sum_{c_g,\mathbf{q},a_1,a_2,\mu} \pi(c_g, \mathbf{q}, a_1, a_2, \mu) W(q_1 + q_2 - c_g) \geq \bar{W}, \tag{14}$$

$$\begin{aligned} \forall \mu, \quad & \sum_{c_g,\mathbf{q}} \pi(c_g, \mathbf{q}, a_1, a_2, \mu) \sum_i \mu_i (U(c_i(c_g, \mu)) - V(e_i(e_g, \mu))) \\ & \geq \sum_{c_g,\mathbf{q}} \pi(c_g, \mathbf{q}, a_1, a_2, \mu) \frac{p(\mathbf{q}|\hat{a}_1, \hat{a}_2)}{p(\mathbf{q}|a_1, a_2)} \sum_i \mu_i (U(c_i(c_g, \mu)) - V(e_i(\hat{e}_g, \mu))) \\ & \forall a_1, a_2, \hat{a}_1, \hat{a}_2, \text{, and where } \hat{e}_g = \hat{a}_1 + \hat{a}_2, \end{aligned} \tag{15}$$

$$\forall \bar{a}_1,\ \bar{a}_2,\ \bar{\mathbf{q}},\ \bar{\mu},\ \sum_c \pi(c_g, \bar{\mathbf{q}}, \bar{a}_1, \bar{a}_2, \bar{\mu}) = p(\bar{\mathbf{q}}|\bar{a}_1, \bar{a}_2) \sum_{c,\mathbf{q}} \pi(c_g, \mathbf{q}, \bar{a}_1, \bar{a}_2, \bar{\mu}), \tag{16}$$

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} \pi(c_g, \mathbf{q}, a_1, a_2, \mu) = 1, \text{ and } \forall c_g,\ \mathbf{q}, a_1, a_2, \mu, \quad \pi(c_g, \mathbf{q}, a_1, a_2, \mu) \geq 0. \tag{17}$$

The main difference between this group regime, Program 3 and the relative performance regime, Program 1, is the set of incentive constraints, namely equations (15) in Program 3 and equations (2) and (3) in Program 1.[2] Constraints (15) ensure that *given* the group's contract with the outsider, there is no alternative action pair Pareto superior for the group. The ability of the group to redistribute consumption and efforts is already incorporated in the internal distribution rules. The incentive constraint represents a mongrel consumer, in which the within-group Pareto weights, μ, are an essential parameter. The within-group weight is a choice variable and need not equal the planner's weights λ. The potential difference between these values will be discussed later.

Now suppose we are under the second technology specification, that each agent works only his own plots. Individual efforts no longer solely depend on aggregate effort and the internal distribution rule, μ, but the simplification used for consumption still applies and the programs are quite similar. This program is labeled Program 4 but it is not written out.

At this point we pause to consider the statistical implications of the predicted within-group allocation of resources. We then return to the analysis of the contract between the group and the outsider.

As is already apparent in the arguments leading up to Programs 3 and 4, there will be an optimal allocation of consumption risk within the group. The μ weights used in Program S will determine levels of consumption; roughly, the higher the weight μ_1 relative to μ_2, the higher will be consumption of agent 1 relative to agent 2. Of course, consumption of agents 1 and 2 will also move around with outputs q_1 and q_2 but only in so far as aggregate consumption $c_g(q_1, q_2)$ moves. Holding aggregate, group consumption constant, variations in q_1 will not influence c_1, nor will variations in q_2 influence c_2.

This sole dependence of consumption on group consumption implies that farmers within the group should pass the econometric tests for full insurance as in Cochrane (1991), Deaton (1993), Mace (1991), and Townsend (1994), but only with respect to group consumption. These tests regress individual consumption on aggregate consumption, often at the village or national level,

[2] In fact, if the previous substitutions were not made then Program 3 would be identical to Program 1 except for the incentive constraints. Everything else, the grid space, the technology constraints, the participation constraint, and the probability measure constraints are unchanged.

to see if the functional relationship in equation (12) holds. This section provides a theory of the level of aggregation at which the tests should be performed. The theory also implies that careful measurement of group membership is necessary. The tests are valid for either technological specification.

Similar arguments can be used to devise tests for production efficiency as in Benjamin (1992), and Feder, et. al. (1991). The test would require calculating the implicit production function, as was done in equation (9), and then comparing marginal products, from the group's view, on each plot. Of course, these production efficiency tests are not relevant under the second technological specification, because farmer's can not allocate their labor across plots.

Though allocations within groups have nice neoclassical properties, there remains an incentive problem in both Programs 3 and 4, of the group relative to the outsider. The group represents a mongrel consumer which must be given an incentive to work hard and in specified amounts across the two plots. This is evident in the incentive constraints. Some numerical examples from Program 4 are revealing of the importance of within-group allocations for the group's incentive problem.

Numerical Example 1

Suppose preferences for farmer i are $U(c_i, e_i) = c_i^{.5}/.5 - e_i^{.5}/.5$ for $i = 1, 2$. Recall again that in Program 4, agent 1 only works plot 1 and agent 2 only works plot 2. Consumption for each farmer is gridded over the range 0 to 25 at intervals of one. Actions on each plot can take on a low value of a_l or a high value of a_h. Similarly, output on either plot can take on a low value of q_l or high value of q_h. In this example, the following numbers were chosen for efforts and outputs: $a_l = 2$, $a_h = 6$, $q_l = 2$, and $q_h = 20$.

Output returns are not correlated. On both plots the probability distribution of output, $p(q|a)$ is described by

	q_l	q_h
a_l	.70	.30
a_h	.30	.70

Finally, the principal is assumed to be risk neutral, and the level, $\bar{W}$, of his participation constraint is set equal to 20 units.

Figure 1 shows expected group consumption, $E(c_g(q_1, q_2))$, as a function of output on plots 1 and 2. Expected consumption is shown only because actual solutions delivered lotteries over adjacent consumption grid points. The planners weight λ_1 of agent 1 is on the x-axis. The weight of agent 2 is not shown, but it is just $\lambda_2 = 1 - \lambda_1$. There are four possible outputs, reflecting all different combinations of high or low output on the two plots. The legend describes which line corresponds to which output. For example, output of "hi,lo" means $q_1 = 20$, and $q_2 = 2$. The solution to the model for all of the λ values has both farmers working the high effort.

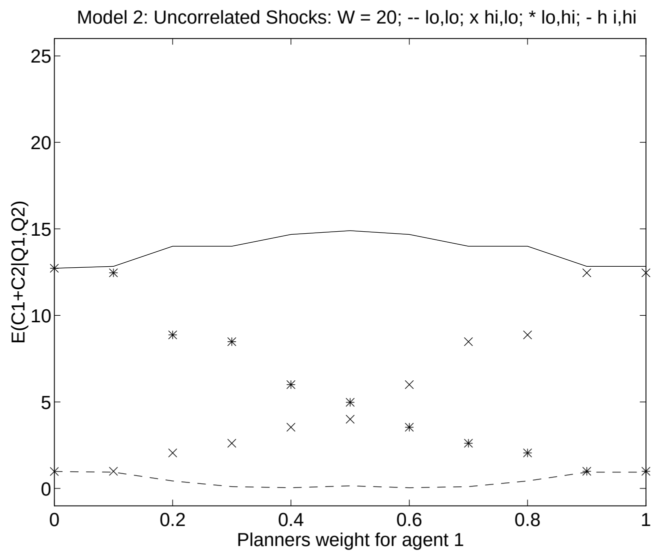

Figure 1: Total group consumption as a function of plot outputs

Starting from the left side of the graph, where $\lambda_1 = 0$, agent 1 receives virtually no weight within the group. (Again, continue to ignore momentarily the potential difference between λ in the master program and the within-group weights, μ.) Not shown in the graph is individual consumption. When $\lambda_1 = 0$, agent 1's consumption is zero for all outputs, while agent 2 consumes all of the group's consumption. Both are assigned high labor effort. Low consumption and high effort of agent 1 has little consequence for group utility because farmer one has low weight. Essentially, agent 1 is a serf. Internally, having agreed on the distribution of wealth, due to skewed ownership or other considerations, agent 1 abides by the agreement, which is to work hard and consume nothing. Notice that the $c_g(\mathbf{q})$ lines for "hi,hi", and for "lo,hi" coincide in figure 1 as do the "lo,lo" and "hi,lo" lines when $\lambda_1 = 0$. Externally, group consumption does not depend on output from the plot farmed by agent one. Internal monitoring and perfect commitment takes care of the potential incentive problem, thus the risk-neutral principal provides full insurance on plot one. In contrast, agent 2, the lord of the group, has a high λ_2 weight. Consequently, the mongrel consumer cares a lot about his effort and the group must be given incentives to make the lord work hard. Hence, the group's consumption and the lord's consumption vary positively with the output of agent 2 on plot two.

This logic prevails more generally, as we let λ_1 increase toward the symmetric weight .5. Over this range group consumption depends primarily on output of agent 2 from plot 2, though output of agent 1 from plot 1 becomes increasingly important. Group consumption, c_g, is ordered with plot outputs as follows: $c_g(q_1=2, q_2=2) < c_g(q_1=20, q_2=2) < c_g(q_1=2, q_2=20) < c_g(q_1=20, q_2=20)$. The ordering reflects the relative importance of each agent's efforts as described above.

At lower levels of the principal's utility, when he has relatively low wealth relative to the group, it is not optimal for both agents to work hard. At $\lambda_1 = 0$ agent 1 still works hard and consumes virtually nothing. There is still no incentive issue there. But at $\lambda_1 = 0$, agent 2 now works the low amount so there is no need to give agent 2 an incentive to work hard. The allocation in this extreme case is equivalent to the allocation for the full-information version of the model. Neither individual nor group consumption are a function of plot outputs. However, as λ_1 increases, the group cares more about agents 1's welfare and is more tempted to have him shirk. Thus, stronger incentives are introduced gradually to induce the group to force agent 1 to work hard. As this happens group consumption will vary with output q_1, because of the need for incentives.

We come back to the potential divergence between λ weights in Program 4 and μ weights in group incentive constrain. In particular, we discuss this issue in terms of the example generating figure 1. When $\lambda_1 = .10$, the optimal μ_1 is equal to .02. In this example, the within-group distribution of weights is more unequal than in the master program. Had we constrained the μ-weights to equal the higher λ-weights, agent 2 would have been subjected to a lottery occasionally assigning lower effort. As noted, in the actual unconstrained solution the relatively wealthy agent 2 is being given an incentive to work hard. One possible intuition is that for unequal μ weights the mongrel consumer is more risk averse making it easier to give incentives to work hard. This allocation would introduce less of a distortion from the benchmark, full-information program, despite the negative effect on the objective function resulting from the group using μ weights rather than λ weights to allocate resources within the group.

Exceptions to this intuition can be found. For example, at $\lambda_1 = .40$ parameter μ_1 moves in the opposite direction, to .42. In this parameterization recommended actions do not change, so we can focus on group consumption. In the constrained $\mu_1 = \lambda_1$ allocation the variance of group consumption is 33.2101 while in the unconstrained $\mu_1 \neq \lambda_1$ allocation the variance is 31.4626. Here allowing the within-group allocation to be more equal reduces the variability of group consumption.

For certain classes of preferences and technologies there can be no divergence between μ-weights and λ-weights. To see this please return to Program 3, in which labor is transferable across plots. Assume sets of feasible consumptions and efforts are continuums just as we did in our analysis of internal group sharing rules but retain the CRRA preference specification of the example just given. These preferences aggregate in the sense of Gorman (1954). In particular equations (12) and (13) reduce to

$$c_i = \frac{\mu_i^{1/(1-\alpha)}}{\mu_1^{1/(1-\alpha)} + \mu_2^{1/(1-\alpha)}} c_g, \tag{18}$$

$$e_i = \frac{\mu_i^{1/(1-\alpha)}}{\mu_1^{1/(1-\alpha)} + \mu_2^{1/(1-\alpha)}} e_g, \tag{19}$$

for $i = 1, 2$. Individual allocations of consumption and effort are linear in group aggregates.

If we substitute these expressions back into the weighted utility function of the incentive constraint (15) in Program 3, a common constant, $k(\mu)$, can be pulled out from both sides of the equation, leaving only a utility function expressed in aggregates c_g and e_g. Similarly in the objective function of Program 3 a scalar, $k(\lambda)$, can be pulled out so the objective function consists of that scalar multiplied by a function of the aggregates, c_g and e_g. The constant makes no difference to the maximization problem, that is, to the information-constrained optimal choice of c_g and e_g, though it still effects the internal distribution of consumption and effort. Varying the weights within the group will not affect this schedule of payments to the outsider in any way!

In such cases the weights μ and λ internal to the group disappear from the problem altogether. It is as if the outsider were facing a single agent who has the choice of effort over the two plots. The consumption allocation to this single agent, really to the group, is determined as in the well-understood, classic, principal-agent model.

The reason that CRRA preferences have this feature is that income expansion paths are linear. Equivalently, the distribution of planner's weights does not effect the equilibrium marginal rates of substitution between effort and consumption. Consequently, all types have the same preference ordering over aggregate consumptions and aggregate effort. If the group voted on the choice of aggregate effort, for example, there would be unanimity. For further discussion of the Gorman class of preferences, of which CRRA is one type, see Gorman (1954) or the exposition in Townsend (1993).

Normally, when agents can not work the plots of each other, this logic fails. Preferences would not aggregate in the example economy under consideration in Program 4. However, special preferences specifications provide exceptions. For example, Holmström and Milgrom (1990) use transferable utility, which not surprisingly, can be aggregated across plots.

2.3 Comparison of Regimes

Finally, we return to the central issue of the paper: Why are there firms? Here that question is posed more precisely: When does the group model dominate the relative performance model? There is a small literature, consisting of Holmström and Milgrom (1990), and Ramakrishnan and Thakor (1991), and Itoh (1993), which considers this problem for variations of the models described. In particular, they find that the correlation of plot returns is a crucial variable in comparing the two regimes. They find that relative performance is better when plot returns are correlated while groups are better when plot returns are uncorrelated. We have already noted the power of the relative performance model in the examples of section 2.1. This paper's main contribution to the literature is to illustrate that the distribution of wealth can be important in determining which regime is preferable. The key difference in our approach is that we do not use transferable utility, unlike in Holmström and Milgrom (1990), nor do we restrict ourselves to symmetric equilibrium as in Itoh (1993) and Ramakrishnan and Thakor (1991).

When outputs across the two plots are correlated, then the relative performance model can dominate, sometimes. To see this please consider a comparison of Programs 2 and 4, the relative performance and group regimes when the technology is such that agents must work their own plots. Consider the following example.

Numerical Example 2

Imagine that our economy is identical to the one described earlier in example 1 except that plot outputs are correlated as shown in table 1. It is easy enough to verify that one agent's effort does not effect the unconditional probability of output on the other's plot.

	$(q_1,\ q_2)$			
(a_1, a_2)	(q_l, q_l)	(q_h, q_l)	(q_l, q_h)	(q_h, q_h)
(q_l, q_l)	.6979	.0021	.0021	.2979
(q_h, q_l)	.2991	.4009	.0009	.2991
(q_l, q_h)	.2991	.0009	.4009	.2991
(q_h, q_h)	.2979	.0021	.0021	.6979

Table 1: Production technology, $p(q_1, q_2|a_1, a_2)$ with correlated returns.

The key factor for the comparison in the correlated case is the distribution of wealth within the group. Figure 2 shows a slice of the three-dimensional Pareto frontier for both models, when the principal is constrained to receive $\bar{W} = 20$ utils. The solid line shows the frontier for the group model, Program 4, and the hyphenated line shows the frontier for the relative performance model, Program 2.

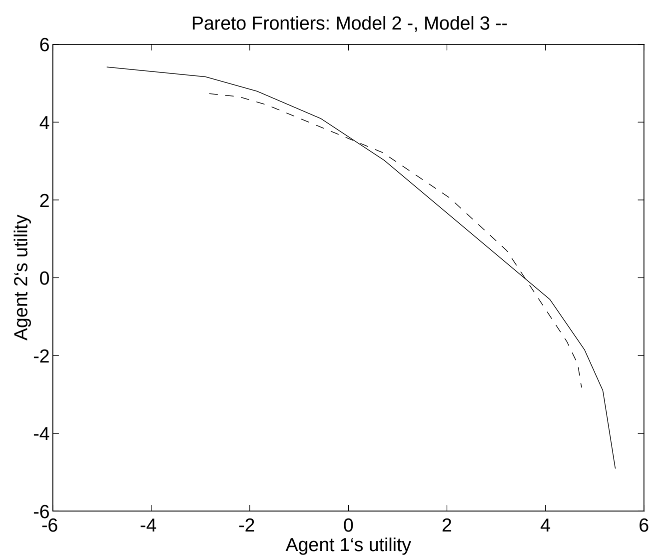

Figure 2: Pareto frontiers for relative performance and group regimes in an example with correlated plot returns.

The relative performance model is Pareto superior to the group model when then is a more egalitarian distribution of the planner's weights, λ_1 and λ_2. However, at extreme distributions of wealth within the group this ranking changes, so that the group model is Pareto superior. Indeed there are utility values not feasible for either agent in the relative performance model. The reason for this is the ability of the group to force a high effort and low consumption allocation on the agents. Such extreme allocations do not satisfy the individual-based incentive constraints in the relative performance program where an individual controls his own effort.

There is a final interesting feature related to the distribution of wealth. There are high levels of the outsider's utility, $\bar{W}$, which can be obtained in the group model but not in the relative performance model. The reason is that in the group model one agent within the group can be used as a "serf" by the group even if the "lord" within the group gets little. The output on both plots is then transferred to the outsider. Such a scheme is not feasible in the relative performance model because, again, the individual incentive constraints require some compensation to each agent to induce high efforts. At high levels of utility, $\bar{W}$, the outsider is using the dominant agent in the group to extract a greater surplus from the group as a whole.

2.4 Decentralization

Like most models with private information, Programs 3 and 4 analyzed groups in isolation, with only a resource constraint to represent their interactions with a principal as a proxy for the rest of the world. Such a simplification is unappealing for several reasons. First, resources are allocated across groups by markets. Neither the cropping groups in Aurepalle nor the partnerships nor the extended Thai kinship networks are isolated from the rest of the world. Second, the principal-agent framework provides no guidance on the extent of the deficit or surplus of the resource constraint, or equivalently, what the level of a risk-neutral principal's utility, $\bar{W}$, should be. As the earlier analysis demonstrated, the level of $\bar{W}$ can strongly affect allocations.

We address these questions by showing that resources can be allocated across groups in a competitive equilibrium, despite private information concerning what is going on within groups and despite collusive behavior within groups relative to outsiders. As before, the groups can enforce within-group allocations while keeping their information private from outsiders. Now, however, instead of dealing with a principal, the members of a group trade as a unit with other groups in a market. The objects traded by the groups are group incentive-compatible insurance contracts, consisting of of transfers to the group conditioned on the recommended internal-weights μ, recommended efforts and observed outputs. Standard results of general equilibrium analysis, including existence and the Welfare Theorems are shown to hold in this environment. The theory shows not only that resources can be allocated by a price system but it establishes the crucial link between endowments and final allocations which is missing in much of agency theory.

A secondary focus of this section is to demonstrate that allocations *within* a group can also be allocated by a price system. Simultaneously with the operation of the across-group market, resources are allocated within groups in a smaller market, naturally one in which participation is limited to members of that group. The result is dual pricing systems, one across groups and one within each group. We discuss at the end of this section the relationship between the two.

Crucial to decentralizing the group insurance market is our choice of a commodity space and consumption sets. Let C_g be the grid of group consumptions, let A_i, $i = 1, 2$ be the grids of total efforts on plot i, and let M be the grid of internal-group Pareto weights, μ. Define $P = C_g \times \mathbf{Q} \times A_1 \times A_2 \times M$. Expanding on E.C. Prescott and Townsend (1994a) and (1994b), the underlying commodity space will be $\Re^n$, where n be the number of points in P. For our private-information economies, modified to handle a continuum of groups, our consumption set is the set of probability measures over P such that the technology and group incentive-compatibility constraints hold. The group's "consumer" problem will be then to choose among probability distributions over the grid, just as in the programs discussed earlier, but now subject to a budget constraint in which these distributions are priced. The resulting economy retains all of the standard features of general equilibrium theory, including convexity of consumption and production sets. Standard results, in particular the Welfare Theorems, will be valid for our economy.

We begin with the group model of the previous section but imagine there is a continuum of groups but only a finite number of group types. Each group type, g, is a fraction $\alpha_g > 0$ of the population. As before, each group is characterized by two agents, their utility functions $U_{gi}(c_i) - V_{gi}(e_i)$, $i = 1, 2$, and by the technology $p(\mathbf{q}|a_1, a_2)$ on the group's two plots of land. The Pareto weights for all group members of a given type g are the same and denoted $\lambda_g = (\lambda_{g1}, \lambda_{g2})$ but weights are allowed to vary across group types.

Any Pareto optimum – in which all groups of the same type are treated ex ante identically – can be found by maximizing a weighted sum of group utilities subject to incentive constraints for actions on the each group's two plots, technology constraints, measure constraints, and an economy-wide resource constraint. Specifically, for assigned Pareto weights λ_g maximize by choice of $\pi_g(c_g, \mathbf{q}, a_1, a_2)$, with $a_1 + a_2 = e_g$, for each group type g, the following program:

Program 5

$$\max_{\pi_g(\cdot)} \sum_g \sum_{c_g, \mathbf{q}, a_1, a_2, \mu} \pi_g(c_g, \mathbf{q}, a_1, a_2, \mu) \sum_i \lambda_{gi} [U_{gi}(c_{gi}(c_g, \mu)) - V_{gi}(e_{gi}(e_g, \mu))]$$

$$\text{s.t.} \sum_g \alpha_g \sum_{c_g, \mathbf{q}, a_1, a_2, \mu} \pi_g(c_g, \mathbf{q}, a_1, a_2, \mu)(q_1 + q_2 - c_g) = 0, \tag{20}$$

$$\begin{aligned} \forall g, \mu \ & \sum_{c_g, \mathbf{q}} \pi_g(c_g, \mathbf{q}, a_1, a_2, \mu) \sum_i \mu_i (U_{gi}(c_{gi}(c_g, \mu)) - V_{gi}(e_{gi}(e_g, \mu))) \\ & \geq \sum_{c_g, \mathbf{q}} \pi_g(c_g, \mathbf{q}, a_1, a_2, \mu) \frac{p(\mathbf{q}|\hat{a}_1, \hat{a}_2)}{p(\mathbf{q}|a_1, a_2)} \sum_i \mu_i (U_{gi}(c_{gi}(c_g, \mu)) - V_{gi}(e_{gi}(\hat{e}_g, \mu))), \\ & \forall a_1, a_2, \hat{a}_1, \hat{a}_2, \text{ and where } \ \hat{e}_g = \hat{a}_1 + \hat{a}_2, \end{aligned} \tag{21}$$

$$\forall g, \ \bar{a}_1, \ \bar{a}_2, \ \bar{\mathbf{q}}, \ \bar{\mu}, \ \sum_c \pi_g(c_g, \bar{\mathbf{q}}, \bar{a}_1, \bar{a}_2, \bar{\mu}) = p(\bar{\mathbf{q}}|\bar{a}_1, \bar{a}_2) \sum_{c_g, \mathbf{q}} \pi_g(c_g, \mathbf{q}, \bar{a}_1, \bar{a}_2, \bar{\mu}), \tag{22}$$

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} \pi_g(c_g, \mathbf{q}, a_1, a_2, \mu) = 1, \text{ and } \ \forall g, c_g, \mathbf{q}, a_1, a_2, \mu, \ \ \pi_g(c_g, \mathbf{q}, a_1, a_2, \mu) \geq 0. \tag{23}$$

Program 5 differs only from Program 3 in the addition of a continuum of heterogenous groups. One constraint altered by switching to a continuum is equation (20), the resource constraint. It allows goods to be transferred across group types and is written in terms of expectation, since with a continuum of groups $\pi_g(c_g, \mathbf{q}, a_1, a_2, \mu)$ is the fraction of that type who will receive the allocation $(c_g, \mathbf{q}, a_1, a_2, \mu)$ in equilibrium. Equations (21) are the group-incentive constraints, equations (22) ensures consistency in allocations with the underlying technology, and (23) are the probability measure constraints. As with Program 3, the above program can be modified to incorporate the technology where each agent may only work a single plot. For that specification, a different grid space needs to be used since individual efforts are no longer solely a factor of aggregate effort e_g and internal weight μ. Nevertheless, none of the decentralization results which follow will be affected by the differences between the two labor substitutability specifications so for simplicity we assume that all groups have the ability to substitute workers across plots.

Our discussion of decentralization is broken into two separate parts. The first part studies decentralization of allocations across groups. By this we mean the group acts as a single entity, purchasing allocations for all of its members. A competitive equilibrium is defined, shown to exist, and the Welfare Theorems are shown to hold. Finally, it is demonstrated that under some conditions within-group allocations can also be supported by a different price system applying only to members within a group.

2.4.1 Across-group decentralization

Following E.C. Prescott and Townsend (1984a) and (1984b), the commodity space is $\Re^n$, where n is the number of points in P, the set of feasible grid points. The consumption set, X_g, for a type g group is an allocation $x_g(c_g, \mathbf{q}, a_1, a_2, \mu) \in \Re^n$ which is a probability measure,

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} x_g(c_g, \mathbf{q}, a_1, a_2, \mu) = 1, \quad \forall c_g, \mathbf{q}, a_1, a_2, \mu, \quad x_g(c_g, \mathbf{q}, a_1, a_2, \mu) \geq 0,$$

group incentive-compatible,

$$\begin{aligned}
\forall g, \mu, \quad & \sum_{c_g, \mathbf{q}} x_g(c_g, \mathbf{q}, a_1, a_2, \mu) \sum_i \mu_i (U_{gi}(c_{gi}(c_g, \mu)) - V_{gi}(e_{gi}(e_g, \mu))) \\
& \geq \sum_{c_g, \mathbf{q}} x_g(c_g, \mathbf{q}, a_1, a_2, \mu) \frac{p(\mathbf{q}|\hat{a}_1, \hat{a}_2)}{p(\mathbf{q}|a_1, a_2)} \sum_i \mu_i (U_{gi}(c_{gi}(c_g, \mu)) - V_{gi}(e_{gi}(\hat{e}_g, \mu))) \\
& \forall a_1, a_2, \hat{a}_1, \hat{a}_2, \text{ and where } \hat{e}_g = \hat{a}_1 + \hat{a}_2,
\end{aligned}$$

and technologically feasible,

$$\begin{aligned}
& \forall g, \ \bar{a}_1, \ \bar{a}_2, \ \bar{\mathbf{q}}, \ \bar{\mu}, \\
& \quad \sum_c x_g(c_g, \bar{\mathbf{q}}, \bar{a}_1, \bar{a}_2, \bar{\mu}) = p(\bar{\mathbf{q}}|\bar{a}_1, \bar{a}_2) \sum_{c_g, \mathbf{q}} x_g(c_g, \mathbf{q}, \bar{a}_1, \bar{a}_2, \bar{\mu}).
\end{aligned}$$

Each group has an endowment, $\xi_g(c_g, \mathbf{q}, a_1, a_2, \mu) \in \Re_n$. In an economy without wealth redistribution the natural endowment is the lottery ξ_g putting mass one on the autarky point $(0, \mathbf{0}, 0, 0, \mu)$, where μ is an arbitrary within-group weight. This endowment point entails no work, no output, and no consumption with certainty. As will be discussed later, in an economy with wealth redistribution, $\xi_g(c_g, \mathbf{q}, a_1, a_2, \mu)$ could be a solution to Program 5.

The price system in this economy is a vector $p(c_g, \mathbf{q}, a_1, a_2, \mu) \in \Re^n$ which maps any allocation into $\Re$. Thus, the budget constraint of any group g is

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} x_g(c_g, \mathbf{q}, a_1, a_2, \mu) p(c_g, \mathbf{q}, a_1, a_2, \mu) \leq \sum_{c_g, \mathbf{q}, a_1, a_2, \mu} \xi_g(c_g, \mathbf{q}, a_1, a_2, \mu) p(c_g, \mathbf{q}, a_1, a_2, \mu). \quad (24)$$

The "consumer's" problem for a group of type g is then to choose $x_g(c_g, \mathbf{q}, a_1, a_2, \mu) \in X_g$ to maximize the objective function

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} x_g(c_g, \mathbf{q}, a_1, a_2, \mu) \sum_i \lambda_{gi} [U_{gi}(c_{gi}(c_g, \mu)) - V_{gi}(e_{gi}(e_g, \mu))],$$

subject to the budget constraint (24).

As in standard theory there are profit-maximizing intermediaries in the economy which take the price system $p(c_g, \mathbf{q}, a_1, a_2, \mu)$ as given and make commitments to supply groups with the underlying commodity. As we will soon see there is constant returns to scale in this technology so without loss of generality we assume that there is only one intermediary.[3]

Let $y(c_g, \mathbf{q}, a_1, a_2, \mu) \in \Re^n$ be the number of units of each commodity point supplied by an intermediary. The intermediary may not commit to give out more consumption than it receives so its production set, Y, is

$$Y = \{y(c_g, \mathbf{q}, a_1, a_2, \mu) \in \Re^n | \sum_{c_g, \mathbf{q}, a_1, a_2, \mu} y(c_g, \mathbf{q}, a_1, a_2, \mu)(c_g - q_1 - q_2) \leq 0\}. \tag{25}$$

The intermediary's problem is to maximize profits,

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} y(c_g, \mathbf{q}, a_1, a_2, \mu) p(c_g, \mathbf{q}, a_1, a_2, \mu),$$

subject to $y(c_g, \mathbf{q}, a_1, a_2, \mu) \in Y$.

The constant returns to scale of Y delivers an equilibrium price system $p^*(c_g, \mathbf{q}, a_1, a_2, \mu)$ proportional to $c_g - q_1 - q_2$ or otherwise there would be no solution to the intermediary's problem. Thus, in equilibrium, consumption c_g "produced" by the intermediary receive a positive price and outputs $\mathbf{q}$ "purchased" by the intermediary receive a negative price. Other elements in the commodity space, quantities of efforts and assigned μ-weights, receive a zero price.

Market clearing requires that $\forall c_g, \mathbf{q}, a_1, a_2, \mu \in P$,

$$\sum_g \alpha_g x_g(c_g, \mathbf{q}, a_1, a_2, \mu) = \sum_g \alpha_g \xi_g(c_g, \mathbf{q}, a_1, a_2, \mu) + y(c_g, \mathbf{q}, a_1, a_2, \mu).$$

With the group's problem, intermediary's problem, and market clearing conditions in place, we can now define a competitive equilibrium.

Definition 1 *A competitive equilibrium is an allocation x_g^* for each group type g, an allocation y^* for the intermediary and a price system p^* such that*

1. *x_g^* solves the group's problem given p^*,*

2. *y^* solves the intermediary's problem given p^*, and*

3. *markets clear.*

[3] To avoid confusion with the firms in the title of the paper, we avoid the standard general equilibrium usage of the term firm and instead call it an intermediary. Besides, in this economy the term intermediary is more descriptive than a firm since the traded objects are insurance contracts and physical production is done by groups when they choose $x_g \in X_g$.

Competitive equilibria exist in this economy as the next theorem shows.

Theorem 2 (Existence) *For the natural endowment point, $\xi_g(c_g, \mathbf{q}, a_1, a_2, \mu) = (0, \mathbf{0}, 0, 0, \mu)$, that is without any initial wealth redistribution, there exists a competitive equilibrium.*

Proof: Recalling that equilibrium prices are linear in $c_g - q_1 - q_2$, each group's income, I_g, is

$$I_g = \sum_{c_g, \mathbf{q}, a_1, a_2, \mu} \xi_g(c_g, \mathbf{q}, a_1, a_2, \mu)(c_g - q_1 - q_2).$$

Under our choice of initial endowments, income for each and every group is zero! Each group's consumer problem is to choose an allocation $x_g \in X_g$ which maximizes utility subject to the budget constraint

$$\sum_{c_g, \mathbf{q}, a_1, a_2, \mu} x_g(c_g, \mathbf{q}, a_1, a_2, \mu)(c_g - q_1 - q_2) \leq I_g = 0.$$

The budget constraint states that the expected net inflow of resources to any group g can not be positive.

In effect, each group type g purchases an insurance contract which is actuarially fair. There is not full insurance for any given group; aggregate c_g moves with $\mathbf{q}$, as required by the incentive constraints. But because there is a continuum of groups of type g, $x_g(c_g, \mathbf{q}, a_1, a_2, \mu)$ represents the fraction of groups of type g in the population experiencing outcome-assignments $(c_g, \mathbf{q}, a_1, a_2, \mu)$. Since there is no uncertainty about aggregates, resource flows sum to zero within each group type g. That is, there is no net flow of resources from groups of one type to another. This implies that each group-type achieves feasible, incentive-compatible insurance without interacting with different group types, just as if each group type was dealing with their own risk-neutral principal with a reservation utility level of zero.

Let $[x_g^*]$ be the vector across group types of solutions to group g consumer problems and set $y^* = \sum_g \alpha_g (x_g^* - \xi_g)$. Substitution of the market clearing conditions into the intermediary's resource constraint verifies that y^* is feasible for the intermediary. Finally, since $p^* y^* = 0$ and $p^* y \leq 0$, $\forall y \in Y$, y^* solves the intermediary's problem. **Q.E.D.**

The proof demonstrates that there is a mapping between a group's budget problem and the principal-agent problem with a risk-neutral principal. The mapping suggests that competitive equilibria are Pareto optimal, which the following theorem verifies.

Theorem 3 (First Welfare Theorem) *Every competitive equilibrium allocation* $[x_g^*]$ *is Pareto optimal.*

Proof: By definition of a competitive equilibrium, the vector across groups $[x_g^*]$ solves the consumer problem for each group type g. Now let $\pi_g = x_g^*$, $\forall g$.

Consider the global optimum problem, Program 5. Because of the continuums, the only connection groups have to one another is through the resource constraint. Thus, Program 5 decomposes into a series of subprograms each characterized by the level of the resource constraint associated with each group type. The allocation $\pi_g = x_g^*$ solves each subproblem at a zero resource level. It is also feasible since Program 5's resource constraint is the sum of the resource constraints of each subproblem. Therefore, $[x_g^*]$ is Pareto optimal. **Q.E.D**

The Second Welfare Theorem follows naturally.

Theorem 4 (Second Welfare Theorem) *Let* $[\pi_g^*]$ *be a Pareto optimum. Then there exists a competitive equilibrium for which* $x_g^* = \pi_g^*$, $\forall g$.

Proof: By the discussion above we can trace out all the Pareto optima across groups by varying the levels of the subprogram resource constraints in Program 5, group by group, in such a way as not to violate the economy-wide resource constraint. Each one of these subproblems is equivalent to the consumer's problem for the corresponding level of income, where for each group g, the Pareto weights λ_g in Program 5, are used in the group g's objective function. It thus suffices to assign these incomes to achieve a competitive equilibrium supporting the desired Pareto optimum. One obvious way to do this is by setting $\xi_g = x_g^*$ and then determining incomes with prices p^*. **Q.E.D.**

The economy's simplicity also provides other interesting implications. First, private information plays only a small role in this economy. It determines the probability distribution of consumption, outputs, efforts, and internal-weights μ, but it has no effect on prices or incomes. The only difference between this economy and one with full information would be the lack of incentive constraints in consumption sets X_g. Prices, resource constraints, and incomes are all unchanged. The interesting differences between the economies will be in efforts, total outputs, and in particular, variability of consumption.

2.4.2 Within-group decentralization

This section demonstrates that in some cases allocations can be decentralized within groups, too. The idea is that simultaneous with the operation of the larger market, there are numerous smaller markets, one per group. Participation in these smaller markets is restricted only to members of that group. The purpose of this section is to demonstrate that such price mechanisms may coexist at different levels within an economy.

The within-group decentralization is much more standard than the across-group decentralization. The allocations which are decentralized are those *conditional* on being recommended internal-weight μ, that is, $\pi_g(c_g, \mathbf{q}, a_1, a_2|\mu)$. To be group incentive compatible the allocation must satisfy the subproblem, Program S, for assigned weight μ. Since in the subproblem, groups may reallocate labor and consumption as they see fit, they care only about the total resource they will receive from other groups. Let $\tau = q_1 + q_2 - c_g$ be the amount of resources paid to the outside. The tax or transfer plus output on the two group plots is the "production" function faced by the group.

This tax is a function of several sources of uncertainty. First, there is randomness in plot returns, $\mathbf{q}$. Second, conditional on output, $\mathbf{q}$, group consumption, c_g, may be random as determined by $\pi_g(c_g|\mathbf{q}, a_1, a_2, \mu)$. Third, as $\pi_g(c_g, \mathbf{q}, a_1, a_2, \mu)$ indicates, τ may depend on recommended actions, too. Given, $\pi_g(c_g, \mathbf{q}, a_1, a_2|\mu)$, uncertainty of plot returns may be summarized by the shocks $(\varepsilon_1, \varepsilon_2)$, where ε_i is the shock to plot i. Uncertainty in the transfer c_g may be described by ε_3, and any randomness in the recommended actions (a_1, a_2) we denote by ε_4.

The tax function can be now be specified in terms of the shocks. Letting $q_i = f_i(a_i, \varepsilon_i)$, the tax function is $\tau(\varepsilon_3, f_1(a_1, \varepsilon_1), f_2(a_2, \varepsilon_2), \varepsilon_4)$. The state of the economy is then $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$, which is drawn from the probability distribution $prob(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$.

Before describing the consumer's and firm's problem in this economy, one explanatory note is required. The underlying commodity points are grids of consumptions and efforts. While this is not a handicap in economies with continuums of agents it is a problem when there are only finite number of agents, as in this economy. The lack of a continuum is a problem because lotteries no longer represent fractions of the population, so the resource constraint requires that a feasible lottery depends on the realization of other agents' lotteries. This can cause several problems.

Consequently, we simply assume that consumptions and efforts are chosen from continuums of sets. This is not a problem for the previous analysis because E.C. Prescott and Townsend (1984b) have shown general equilibrium analysis of private information economies can be done even when the choice variables are probability measures over continuums of sets. One problem is that under this assumption there is a continuum of shocks $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$. Rather than developing notation to handle this we simply assume that there are only a finite number of states $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ which occur in equilibrium. This assumption does not alter the thrust of the following results.

For each group g, let $p(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ be the state-contingent price of consumption and $w(\varepsilon_4)$ be the state-contingent wage. We are assuming that the within-group market opens after the contract with the outsider has been reached, after the internal-weight μ has been assigned, but before actions (a_1, a_2) have been recommended. Let $c_{gi}(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$, $i = 1, 2$, be agent i's state-dependent consumption and $e_{gi}(\varepsilon_4)$, $i = 1, 2$, be agent i's labor supply. Agents receive wage income, profits from the firm, r_i, and any initial transfers, t_i. The consumer's problem for each household i in group g is then

$$\max_{c_{gi}(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4), e_{gi}(\varepsilon_4)} \sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4} prob(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)[U_{gi}(c_{gi}(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)) - V_{gi}(e_{gi}(\varepsilon_4))]$$

subject to a budget constraint

$$\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4} p(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) c_{gi}(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) \leq \sum_{\varepsilon_4} w(\varepsilon_4) e_{gi}(\varepsilon_4) + r_i + t_i.$$

The firm or cooperative also takes internal prices as given. It maximizes profits by purchasing labor from the market, $n(\varepsilon_4)$, allocating it among the two plots, $n_1(\varepsilon_4), n_2(\varepsilon_4)$, and then selling the state-contingent output, $y(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$. It's production function is the tax, $\tau(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ plus plot outputs. The firm's problem is

$$\max_{y(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4), n_1(\varepsilon_4), n_2(\varepsilon_4), n(\varepsilon_4)} \sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4} p(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) y(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) - \sum_{\varepsilon_4} w(\varepsilon_4) n(\varepsilon_4)$$

subject to $\forall \varepsilon_4$, $\sum_i n_i(\varepsilon_4) = n(\varepsilon_4)$, and

$$\forall \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \quad y(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) \leq \sum_i f_i(n_i(\varepsilon_4), \varepsilon_i) - \tau(f_1(n_1(\varepsilon_4), \varepsilon_1), f_2(n_2(\varepsilon_4), \varepsilon_2), \varepsilon_3, \varepsilon_4).$$

Market clearing within the group is

$$\forall \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \quad y(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) = \sum_i c_{gi}(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4),$$

and $\forall \varepsilon_4, \; n(\varepsilon_4) = \sum_i e_{gi}(\varepsilon_4)$.

If the production technology, taxes plut outputs, is convex in $n_i(\varepsilon_4)$, $i = 1, 2$, then a competitive equilibrium is optimal and every internal optimum can be achieved in a decentralized competitive equilibrium, possibly with wealth transfers, t_i. If the technology is not convex, then the standard results do not apply.[4]

The applicability of the within-group analysis depends on the example economy. For example, the Indian cropping groups are relatively small and do not use a price system internally. However, the Thai extended families or the the Aymara Indian villages are larger and might use a decentralized mechanism. Related, many large organizations, such as corporations use internal prices, often called transfer pricing, to allocate resources within the organization.

3 Cropping Groups as Information Monitors

This section analyzes a second class of prototypes which are designed to study an alternative definition of a cropping group or firm. A cropping group in this section is an arrangement under which a set of households monitor each other by jointly farming plots of land. The organizational design question for the planner is to decide which plots are jointly farmed by a group and which are solely farmed by individuals. Before describing in detail what this choice entails, and also to allow for a clear comparison between this and the previous sections' definitions of a firm, it is helpful to first develop the model.

Imagine that are two households and three plots of land. As in sections 2.1 and 2.3 the utility function $U(c_i) - V(e_i)$ describes farmer i's preferences, where e_i is the total amount of effort worked by farmer i, and c_i is farmer i's consumption. Again, we denote $\mathbf{e_{i\bullet}} = (e_{i1}, e_{i2}, e_{i3})$ as the vector of plot-specific efforts by farmer i. The probability of the output vector $\mathbf{q} = (q_1, q_2, q_3)$, across the three plots is described by the function $p(\mathbf{q}|\mathbf{e_{1\bullet}} + \mathbf{e_{2\bullet}})$. The production function can be specified quite generally to incorporate correlated returns and complementarities in production but for simplicity we assume that plot j's return only depends on the total effort on plot j and that plot returns are uncorrelated. Finally, there is an outsider with preferences over the surplus, $q_1 + q_2 + q_3 - c_1 - c_2$. For simplicity we assume he is risk-neutral.

Plots may be allocated in two ways. If plot j is assigned only to farmer i then farmer i alone

[4] Commodity spaces exist which can handle the non-convexities but the ones explored by the authors had little economic content.

can work it. In this case his effort e_{ij} is private information so there is a moral-hazard problem. Alternatively, plot j may be assigned jointly to farmers one and two. In this case they observe each other's efforts so by revelation principle arguments each may be induced to costlessly reveal each other's efforts to the outsider. This argument only works if there is no collusion against the outsider and we assume that this is the case. More will be said on collusion shortly.

To make the choice between joint and individual farming non-trivial two costs are imposed on groups. First, a utility cost is imposed on each agent per jointly farmed plot. The idea is that there are some diseconomies of scale or coordination costs in the number of farmers working together. Second, for consistency with the monitoring focus of this section, agents' efforts must be supplied in equal amounts to each group plot, as if they were monitoring by simultaneously working together.

The lack of collusion is the key difference between this section and the previous one but the two prototypes are not inconsistent. Despite appearances, collusion is not incompatible with land assignments in which there are individually farmed plots. If agents could collude then, as in section 2.4, allocations would need to respect the ability of agents to write their own internal contracts, but now their subproblem itself would include incentive constraints because efforts are not observed on individually operated plots.[5] Such a program can be written down but computing it – let alone analyzing it – is difficult. Consequently, we restrict our focus in this section to the no-collusion case and concentrate on the assignment monitoring feature of the firms.

The prototype in this section is designed to capture certain features of the economies described in the introduction. Most apparent is the connection to the observed land assignments in Aurepalle. In addition to their group obligations, cropping group members often have their own plots which they work separately from the group. On the group plots, cropping group members claim they work together in equal amounts, suggesting that there is monitoring and that this section's equal effort assumption is reasonable. Observed group size is also limited, typically there are only two or three group members and never more than five, suggesting there are diseconomies of scale as assumed here. Another common land assignment in Aurepalle is for landowners to split their land among tenants rather than renting it to groups and for individuals to farm their own plots

[5] One complication in combining the approaches is that allocations may satisfy one definition of a firm and not the other. As we discussed in the introduction our motivation is to model characteristics of some firms. Underlying much of our analysis is that there can be different reasons for firms to exist. This suggests that the firm may be too broad a concept and that it would be beneficial to distinguish firms by their *raison d'etre*.

rather than combining them into one production unit.

By the standard introduced in this section the communal land of the Aymara would not constitute the land of a group unless it is jointly farmed. However, if there were a requirement that all communal plots be farmed at the same time, it would be suggestive of a monitoring component, as assumed here. Similarly, the individual operated plots of the families in Burkina Faso would not constitute group plots, though any plots jointly farmed by males and females would constitute group plots. The prototype of this section also focuses our attention again to extended familiy arrangements in Thailand and to the distinction of whether or not extended families are jointly farming plots or instead only sharing in consumption risk.

Returning to the model, our aim is to jointly solve for land assignments and allocations. Unfortunately, reaching this goal is not straightforward. Complications arise because the grid of feasible consumption, output, and effort vectors depend on the land assignment as does the specification of the incentive constraints. In particular, as the number of cases checked increases – that is, as the number of land assignments increases – so does the size of the linear program. We reduce this problem by computing the optimal allocation for each given land assignment. The optimal land assignment and corresponding allocation will then be the one which maximizes the objective function. Sometimes, however, the linear program designed to compute a given land assignment is still too large for practical purposes. These problems are avoided by assuming a particular specification of the utility function and then using dynamic programming techniques developed in E.S. Prescott (1995) to compute the model. More will be said on our algorithm later.

There are numerous possible land assignments, including those which specify who works which plot. However, since plot j's return is a function of total effort on the plot, $e_{1j} + e_{2j}$, and is independent of other plot returns, it does not matter whether farmer i works plot j, per se. What matters is the number of private and group plots worked by both farmers. Table 2 lists the possible assignments.

Columns one and two list the number of plots worked solely by agents one and two, respectively, while column 3 lists the number of plots jointly farmed. For example, the assignment in the first row $(3, 0, 0)$ means that agent one works all three plots by himself. Row 6, the $(1, 1, 1)$ assignment indicates that each agent works one plot by himself and one plot jointly. This latter assignment is of particular interest because it most closely resembles typical cropping groups in Bolivia or

Number of plots worked by:		
Agent one alone	Agent two alone	The group
3	0	0
2	1	0
1	2	0
0	3	0
2	0	1
1	1	1
0	2	1
1	0	2
0	1	2
0	0	3

Table 2: Feasible land assignments.

Aurepalle.

In the interest of brevity, we describe only the $(1,1,1)$ or cropping group land assignment. In particular, we first describe the domain and the incentive constraints for this problem and then setup the planning problem. Later, we will indicate which part of the program must be modified to handle other land assignments. Further, without loss of generality we assume that for the $(1,1,1)$ land assignment, agent one's plot is plot one, agent two's plot is plot two, and the group's plot is plot three.

Let $P = \mathbf{C} \times \mathbf{Q} \times \mathbf{E_{1\bullet}} \times \mathbf{E_{2\bullet}}$ be the cross-product of the sets of consumption, output, and effort vectors points. Each set, $\mathbf{E_{i\bullet}}$, contains vectors which indicate agent i works zero effort on any, some, or all plots. The assumption is that if agent i is not assigned to plot j then he supplies zero effort on it, that is $e_{ij} = 0$. If he is assigned to plot j then we require that for $\mathbf{e_{i\bullet}}$ to be feasible, $e_{ij} > 0$.

The land assignment, through the aforementioned assumptions, determines the set of feasible grid points. If X is the set of feasible grid points then for the $(1,1,1)$ land assignment

$$X = \{(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) \in P | e_{11} > 0,\ e_{12} = 0,\ e_{21} = 0,\ e_{22} > 0,\ e_{13} = e_{23} > 0\}. \tag{26}$$

The conditions on X require that agent one works positive effort on plot one, does not work plot two, and works an amount on plot three which is not only positive but equal to farmer two's effort. The conditions similarly restrict agent two's feasible grid points. Other restrictions, such as maximum total effort per plot or maximum total effort by an agent, are easily incorporated into X.

Given the land assignment, the choice variable for the program is $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$, the joint probability of consumptions, outputs, and efforts. Aside from the domain, X, only incentive constraints are affected by the land assignment. Consequently, we describe them before writing out the entire program.

For the $(1, 1, 1)$ land division there are two sources of private information, agent one's action on plot one and agent two's action on plot two. Efforts on plot 3 do not require incentive constraints because the plot is jointly farmed. Incentive constraints on agent one are for each e_{13}, e_{11}, the allocation $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$ must satisfy

$$\begin{aligned} &\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})[U(c_1) - V(e_{11}+e_{13})] \\ &\quad \geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) \frac{p(\mathbf{q}|\hat{e}_{11}, e_{22}, e_{13}+e_{23})}{p(\mathbf{q}|e_{11}, e_{22}, e_{13}+e_{23})}[U(c_1) - V(\hat{e}_{11}+e_{13})], \end{aligned} \tag{27}$$

for all feasible $\hat{e}_{11}$. The incentive constraint is designed to stop agent one from deviating on plot one. Since the other agent does not work plot one, total effort on plot one is e_{11}. Both agents' efforts on plot three are publicly observed so agent one must follow the e_{13} recommendation on plot three.

Agent two's incentive constraints are similar taking the form that for each e_{23}, e_{22}, the allocation $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$ must satisfy

$$\begin{aligned} &\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})[U(c_2) - V(e_{22}+e_{23})] \\ &\quad \geq \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{2\bullet}}} \pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) \frac{p(\mathbf{q}|e_{11}, \hat{e}_{22}, e_{13}+e_{23})}{p(\mathbf{q}|e_{11}, e_{22}, e_{13}+e_{23})}[U(c_2) - V(\hat{e}_{22}+e_{23})], \end{aligned} \tag{28}$$

for all feasible $\hat{e}_{22}$.

With the domain and incentive constraints as specified, we can now proceed to the description of the program. Given the land assignment the planner's problem is to maximize the weighted sum of agent's utilities by choosing $\pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$ over the domain X, as defined by (26). Denote $g(y)$ as the disutility accruing to each agent from working y number of group plots; with one group plot, as here, the disutility is $g(1)$. Letting (λ_1, λ_2) be the planner's weights on the agents and $\bar{W}$ the reservation utility of the principal, the program is

Program 6

$$\max_{\pi()} \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})[\lambda_1(U(c_1)-V(e_1))+\lambda_2(U(c_2)-V(e_2))-g(1)]$$

$$\text{s.t.} \quad \sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}})(q_1+q_2+q_3-c_1-c_2) \geq \bar{W}, \tag{29}$$

$$(27), \quad (28),$$

$$\forall \bar{\mathbf{q}}, \bar{\mathbf{e}}_{\mathbf{1\bullet}}, \bar{\mathbf{e}}_{\mathbf{2\bullet}}, \quad \sum_{\mathbf{c}} \pi(\mathbf{c},\bar{\mathbf{q}},\bar{\mathbf{e}}_{\mathbf{1\bullet}},\bar{\mathbf{e}}_{\mathbf{2\bullet}}) = p(\bar{\mathbf{q}}|\bar{\mathbf{e}}_{\mathbf{1\bullet}},\bar{\mathbf{e}}_{\mathbf{2\bullet}}) \sum_{\mathbf{c},\mathbf{q}} \pi(\mathbf{c},\mathbf{q},\bar{\mathbf{e}}_{\mathbf{1\bullet}},\bar{\mathbf{e}}_{\mathbf{2\bullet}}), \tag{30}$$

$$\sum_{\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}} \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) = 1, \quad \pi(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) \geq 0. \tag{31}$$

Program 6 is similar to Program 1, the relative performance model. The term $g(1)$ in the objective function is removed from each agent's portion of the objective function because $\sum_i \lambda_i g(y) = g(y)$. As in earlier sections, equation (29) is the outsider's utility, equations (30) are the constraints which ensure that chosen allocations are consistent with the underlying technology, and constraints (31) ensure that the choice variable is a probability measure. Finally, (27) and (28) are the previously discussed incentive constraints.

Planning problems for other land assignments can be setup by modifying Program 6. In fact, all that needs to be changed is the domain X and the form of the incentive constraints. For example, the $(3,0,0)$ land assignment would require that $X = \{(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) \in P | e_{1j} > 0, \ e_{2j} = 0, \ j = 1,2,3\}$ and that there be incentive constraints for agent one on each and every plot. The $(0,0,3)$ assignment, that is, all plots are jointly farmed, would require $X = \{(\mathbf{c},\mathbf{q},\mathbf{e_{1\bullet}},\mathbf{e_{2\bullet}}) \in P | e_{1j} = e_{2j} > 0, \ j = 1,2,3\}$ and no incentive constraints. In the interest of brevity we do not present the various combinations explicitly.

Our algorithm is to compute the solution to the Program 6 version of each land assignment and then take the best one as measured by the objective function, holding λ_1, λ_2, and $\bar{W}$ fixed. Unfortunately, computation is not practical because of the program's large dimensions. Adding the third plot, as compared with the two-plot, relative-performance model in section 2.1, greatly increases the number of feasible actions along with the number of outputs. Coarse grids for P make computation feasible but the grids are sufficiently coarse to make some results an artifact of the grid. Consequently, an alternative strategy is used. Assumptions are made which allow the problem to be broken into smaller programs and then these (still large) programs are computed using the dynamic programming methods developed in E.S. Prescott (1995). The full details of the algorithm, including a description of the dynamic programming method, are listed in the computational appendix. The following paragraphs provide a summary, which leaves out many details. However, it does indicate that our assumptions on preferences and technology greatly facilitated computation.

Our first simplification is to assume that disutility is linear in effort. This assumption is important because it means that an agent's effort level on a group plot does not affect his incentives to work on his own private plot. The reason is that effort on the group plot lowers an agent's utility by a scalar amount without altering the agent's marginal tradeoff between consumption and effort, the factor which matter for incentive constraints.

A second useful result follows from the outsider's risk-neutrality and the independence of production returns across plots. Using Holmström (1982), these assumptions imply that an agent's compensation depends only on outputs on the plots he works. That is, he should bear only output risk associated with his own plot. Related, consumption levels are a function of output on the privately worked plot and not a function of outputs on group plots.

When combined these assumptions imply that the global problem can be broken into separate subproblems; a subproblem for plots individually worked by agent one, another for plots individually worked by agent two, and a final one for group plots. Once the full range of subproblems is calculated, the optimal combination of them, given the land assignment, is calculated. One more calculation, choosing the land assignment which maximizes the objective function, delivers the optimal land assignment.

3.1 Numerical Results

To illustrate the effect of parameters on the optimal organizational structure, we present two numerical examples. Both examples use the following grid spaces

$$\begin{aligned}
&\mathbf{C} = [0, .02, .04, ..., 1.6] \times [0, .02, .04, ..., 1.6], \\
&\mathbf{Q} = \{0, 1\} \times \{0, 1\} \times \{0, 1\}, \\
&\mathbf{E_{i\bullet}} = \{W \times W \times W | \textstyle\sum_j e_{ij} \leq 12\},
\end{aligned}$$

where W is the set of whole numbers. The $\mathbf{E_{i\bullet}}$ grid contains vectors of non-negative integers such that no more than 12 total units of effort are made by individual i. Letting $P = \mathbf{C} \times \mathbf{Q} \times \mathbf{E_{1\bullet}} \times \mathbf{E_{2\bullet}}$, the domain of grid points is

$$X = \{\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}} \in P |\ \forall j,\ \sum_i e_{ij} \leq 8\},$$

that is, no more than 8 units of effort total may be worked on a plot.

There are two outputs per plot, 0 and 1. Plot returns are independent and only a function of total effort on that plot. The probability distribution of the outputs on each plot, j, is

$$p(q_j = 0 | e_{1j} + e_{2j}) = 1 - ((e_{1j} + e_{2j} - .9)/25)^{.5},$$
$$p(q_j = 1 | e_{1j} + e_{2j}) = ((e_{1j} + e_{2j} - .9)/25)^{.5}.$$

Probabilities of the high output on each plot vary from 0.06, if one unit of effort is supplied, to 0.53, if the maximum feasible amount of effort, eight, is supplied to the plot. The probability distribution of the high output is concave in effort. The joint distribution of the vector of plot outputs is easily calculated.

For each agent i, preferences are described by

$$\frac{c_i^\alpha}{\alpha} - k_e(e_{i1} + e_{i2} + e_{i3}) - k_g y,$$

where y is the number of plots worked by the group, and k_e and k_g are scalars. As discussed earlier, disutility is linear in the amount of total effort $e_i = \sum_j e_{ij}$. The parameters α, k_e, and k_g will be varied in the examples, and so will the distribution of wealth, as determined by the level of the principal's utility $\bar{W}$ and the Pareto weights, λ_1 and λ_2. We do not consider agents which are heterogenous in the preference parameters.

Numerical Example 3

Our first example in this section uses the following parameters:

$$\alpha = .5, \quad k_e = .125, \quad k_g = .075, \quad \bar{W} = 0.$$

Table 3 lists the optimal land assignment as a function of the Pareto weight, $\lambda_1 \in [0, .5]$. At $\lambda_1 = 0$ there is the most inequality, while at $\lambda_1 = .5$ the agents are valued equally. At $\lambda_1 = .5$ all plots are jointly farmed, that is, all are group plots. Essentially, the entire economy is one large

	Number of plots worked by:		
	Agent one	Agent two	The
λ_1	alone	alone	group
0	2	0	1
.1	1	0	2
.2	1	0	2
.3	0	0	3
.4	0	0	3
.5	0	0	3

Table 3: Optimal land assignments in example 3 as a function of the Pareto weights.

group. Not shown are the consumption features of the optimal contract, but not surprisingly agents are fully insured because there are no hidden efforts.

Groups predominate in this example economy because the parameter k_g is relatively small. In particular, at $\lambda_1 = 0$ one plot is still a group plot. In this case, each agent works 3 units of effort on the group plot while agent one works, in addition, 2 units of effort on each of the other two plots. Interestingly, despite the low Pareto weight on agent one, it is not worthwhile to make him do all of the work, as in a $(3, 0, 0)$ assignment. Apparently, as $\lambda_1 \to 0$, his effort level increases (and his consumption levels decrease) and the incentive constraints on individually operated plots become increasingly costly to sustain. Consequently, it becomes worthwhile to keep agent two as a monitor on the third plot, despite the losses to the objective function at $\lambda_2 = 1$, rather than compensating agent one enough to induce him to work the higher effort. This monitoring feature does not always dominate at extreme Pareto weights. For example, at $\bar{W} = -2$, that is, the principal makes a net transfer to the agents, agents' efforts decline, monitoring is no longer required, and the optimal land assignment is $(3, 0, 0)$, that is agent one works all of the plots.

Numerical Example 4

Other parameter values can deliver the cropping group land assignment, $(1, 1, 1)$. Consider the parameterization

$$\alpha = .5, \quad k_e = .1, \quad k_g = .2, \quad \bar{W} = 0.$$

Now there is a much larger utility cost from working a group plot and a slightly lower amount of disutility from effort. Table 4 lists the optimal land assignments as a function of the Pareto

	Number of plots worked by:		
	Agent one	Agent two	The
λ_1	alone	alone	group
0	2	0	1
.1	2	0	1
.2	2	0	1
.3	2	0	1
.4	2	0	1
.5	1	1	1

Table 4: Optimal land assignments in example 4 as a function of the Pareto weights.

weight, λ_1. At $\lambda_1 = .5$, we observe the cropping group assignment $(1, 1, 1)$. The contract looks much as one might expect. Agent's contracts are identical and consumption for each agent is solely a function of output on his privately worked plot. Again, we do not show the levels of consumption, but suffice it to say that each agent receive higher consumption if he produces a higher level of output on his individually-worked plot. As $\lambda_1 \to 0$, agent one quickly becomes the only agent working a plot on his own, and there is no variation in the optimal land assignment.

An interesting parameter to vary is $\bar{W}$, the principal's utility. Table 5 lists the assignments as a function of $\bar{W}$ for the case when $\lambda_1 = .5$, that is, when the two agents are ex ante identical.

At $\bar{W} = 2$, the entire economy is one large group and in particular, both agents are working their maximum effort in order to make large transfers to the principal. The principal is using the groups as a monitoring device. The level of $\bar{W}$ requires so much resource extraction from the agents that the program cannot induce the necessary high efforts from individually worked plots. Consequently, the program gives up on avoiding the utility cost from working group plots and uses the agents as monitors.

As $\bar{W}$ is lowered, the number of group plots declines monotonically. At the lower levels of $\bar{W}$ there appears to be an asymmetry between the efforts of the two agents. This is an artifact of the preferences and the desired effort. At these levels, the principal is recommending that agents assigned to a plot work the minimum amount of effort, which is one. Consequently, there is no incentive problem, since zero effort is not feasible on a plot if you are assigned to work it. Second, disutility is linear in effort so for a given level of effort it does not matter which agent does the work.

$\bar{W}$	Number of plots worked by: Agent one alone	Agent two alone	The group
-3	3	0	0
-2	2	0	1
-1	2	0	1
-.5	1	1	1
0	1	1	1
.5	1	1	1
1	1	0	2
2	0	0	3

Table 5: Optimal land assignments in example 4 as a function of the principal's utility $\bar{W}$.

In addition to the cropping group motivation, our model is applicable to the large literature on land reform. In Bolivia during the 1950's, and in other parts of Latin America at different times, many large estates or Haciendas have been broken up, motivated by the desire to redistribute wealth.[6] The last comparative statics exercise presented, analyzing the effect of the principal's utility, $\bar{W}$, on optimal organizational form is highly relevant to this literature. Our results are consistent with the idea that breaking up large estates is a better way to reallocate wealth from the principal to the agents, than a centralized scheme with communal work and shared consumption. Note, however, that some communal plots can remain in the optimal assignment for intermediate ranges of $\bar{W}$, as maybe with the communal holdings of the Aymara. Interesting exercises would include searching for parameters in which the change in the number of group plots is not monotonic in $\bar{W}$, or studying the robustness of the results under different specifications of group costs.

This section focused on the monitoring aspect of firms. It did this by solving an assignment problem, that is, allowing the planner to choose the location at which agents work. The next section continues the assignment theme but in a different context.

4 Firms as Long-term Relationships

A characteristic of firms, not yet emphasized in the paper, is that they are often long-term arrangements. A firm usually performs a sequence of separate tasks over a period of time. For

[6]Related, there has been extensive collectivization of agriculture by the state in Latin America, and more recently, a tendency to break up these collectivized institutions. See Carter (1987) and Meyer (1989) for discussions of the latter process.

example, the Indian cropping groups, as well as tenants or owner-operators, farm the same plots during an entire cropping season. They plow, sow, and fertilize their plot, to name only a few of the tasks necessary to grow crops.

The alternative to long-term arrangements is to hire inputs on a per task basis from a spot market. Such arrangements are common and coexist with long-term arrangements. For example, there is an active spot market for bullock services in Aurepalle. See Mueller and Prescott (1996) for a description. Similarly, many corporations outsource, that is, hire other companies to perform some tasks.

Our final class of prototypes is designed to study these questions by analyzing long-term versus short-term arrangements in an environment where the length of an arrangement is endogenous. Production takes place in two stages and between the two stages the planner has the option of switching a farmer from his original plot to a new one. Switching affects the information structure in ways which may or may not be beneficial.

In these models, agents who are not switched to new plots, that is they work the same plot of land for the entire season, are considered to be in long-term arrangements. In contrast, agents who are switched to new plots are considered to be in a sequence of short-term arrangements which we interpret loosely as corresponding to participation in a spot market.

4.1 The Role of the Production Function

This section studies the role of the production function in determining optimal assignments of labor. Two specifications of the production technology are studied. In the first, a plot's output is only a function of the sum of interim and final-stage efforts and a random shock. Under this assumption, plus one other on the production function, the optimality of no-switching is demonstrated. Switching is disadvantageous because effort deviations must be prevented on both plots, that is, if agents were switched there are additional incentive constraints relative to the no-switch problem. The argument is analogous to the earlier analysis of the relative performance model with uncorrelated returns and labor substitutability across plots. In that model the question was whether or not to assign each agent to work both plots or to work separate plots.

The second specification of the production technology assumes interim and final stage efforts are complements rather than substitutes. The complementarities are across stages, not across plots. Certain combinations of efforts are imagined to be revealing of the levels of effort. We conjecture that such a technology induces switching because such short-term arrangements prevent the agent from simultaneously altering a plot's inputs at both stages. That is, if the agent was to stay on the same plot he would jointly alter interim and final stage efforts to avoid revealing combinations of efforts. Again, there is an analogy to earlier arguments. In particular, for reasons discussed later, the problem is similar to the comparison between the relative performance and group regimes with correlated plot returns and no labor substitutability across plots.

Imagine there are two agents, $i = 1, 2$, and two plots of land $i = 1, 2$. Agent i starts on plot i. There are two stages to production, an interim stage and a final stage. Agent i's initial effort is denoted a_i and his final effort is denoted b_i. Agents may be switched after the initial stage.

Output on plot i, q_i, is a function solely of efforts worked on that plot and a plot-specific random shock. Shocks across plots are uncorrelated. If agents are not switched, the probability distribution of output on plot i is described by the function $p(q_i|a_i + b_i)$. However, if agents are switched then the production function is written $p(q_i|a_i + b_j)$, $j \neq i$. Agent i's initial effort is always an input onto plot i because he starts there.

Agents' receive utility from consumption and disutility from efforts. Utility is separable and written $U(c_i) - V(a_i) - V(b_i)$, where U and V are strictly concave. The utility function implies that for a given level of total effort, agents prefer to spread their effort across the two stages. As before, feasible consumptions, outputs, and efforts are restricted to finite grids. In particular, we also assume that a_i and b_i are chosen from identical sets which consist of elements spaced equally far apart. We also assume that both agents have equal Pareto weights.

For analytical reasons we make several simplifications. First, we model the switching as a discrete decision made by the planner at time of contracting. Thus, all participants in the economy know initially whether there will be a switch or not. Let $d = 0$ denote no switch and $d = 1$ denote switch.

Our second simplification is to allow only the planner to send messages recommending actions immediately after contracting. This assumption precludes the planner from recommending an interim-stage effort a_i and then after the interim stage effort is taken, sending a random message which recommends final-stage effort b_i.[7] We would rather not delve into the Revelation Principle issues related to this assumption but suffice it to say that this communication structure implies

[7] Examples can be generated where such a strategy is beneficial. Unfortunately, it greatly complicates the analysis of the switching problem.

that each agent is initially recommended a sequence (a_i, b_i) of actions. Finally, we also restrict our focus to symmetric allocations. In this example, symmetry means that agents face the same contract *and* are recommended the same sequence of actions. That is, if agent one is recommended (a_1, b_1) then so is agent two. Essentially, we are restricting ourselves to symmetric, pure strategy, Nash equilibrium in the game played between the two agents. However, realizations of output may still differ across the two plots. These assumptions are admittedly strong but they help with the analytical results.

Since the treatment of the agents is symmetric, we can the write the program in terms of maximizing the utility of agent one since his allocation is identical ex ante to that of agent two. The choice variables, framed in terms of agent 1, are the probability distributions $\pi(c_1, q_1, q_2, b_1, a_1|d)$ and the switch variable $d \in \{0, 1\}$. Output q_2 must be included in π because if agent 1 is switched his final-stage effort, b_2, affects q_2.

Program 7

$$\max_{\pi(\cdot|d), d\in\{0,1\}} \sum \pi(c_1, q_1, q_2, b_1, a_1|d)[U(c_i) - V(a_i) - V(b_i)]$$

$$\text{s.t.} \quad \sum \pi(c_1, q_1, q_2, b_1, a_1|d)(q_1 - c_1) \geq \bar{W}/2, \tag{32}$$

if $d = 0$, then for all assigned a_1, b_1,

$$\begin{aligned} &\sum_{c_1,q_1,q_2} \pi(c_1, q_1, q_2, b_1, a_1|d=0)[U(c_1) - V(a_1) - V(b_1)] \\ &\quad \geq \sum_{c_1,q_1,q_2} \pi(c_1, q_1, q_2, b_1, a_1|d=0)\frac{p(q_1|\hat{a}_1 + \hat{b}_1)}{p(q_1|a_1 + b_1)}[U(c_1) - V(\hat{a}_1) - V(\hat{b}_1)], \quad \forall \hat{a}_1, \hat{b}_1, \end{aligned} \tag{33}$$

if $d = 1$, then for all assigned a_1, b_1,

$$\begin{aligned} &\sum_{c_1,q_1,q_2} \pi(c_1, q_1, q_2, b_1, a_1|d=1)[U(c_1) - V(a_1) - V(b_1)] \\ &\quad \geq \sum_{c_1,q_1,q_2} \pi(c_1, q_1, q_2, b_1, a_1|d=1)\frac{p(q_1|\hat{a}_1 + b_1)}{p(q_1|a_1 + b_1)}\frac{p(q_2|a_1 + \hat{b}_1)}{p(q_2|a_1 + b_1)}[U(c_1) - V(\hat{a}_1) - V(\hat{b}_1)], \ \forall \hat{a}_1, \hat{b}_1 \end{aligned} \tag{34}$$

$\forall d, \bar{q}_1, \bar{q}_2, \bar{b}_1, \bar{a}_1,$

$$\sum_{c_1} \pi(c_1, \bar{q}_1, \bar{q}_2, \bar{b}_1, \bar{a}_1|d) = p(\bar{q}_1|\bar{b}_1, \bar{a}_1)p(\bar{q}_2|\bar{b}_1, \bar{a}_1) \sum_{c_1,q_1,q_2} \pi(c_1, q_1, q_2, \bar{b}_1, \bar{a}_1|d) \tag{35}$$

$$\forall d, \sum_{c_1,q_1,q_2,b_1,a_1} \pi(c_1, q_1, q_2, b_1, a_1|d) = 1, \quad \pi(c_1, q_1, q_2, \bar{b}_1, \bar{a}_1|d) \geq 0. \tag{36}$$

Equation (32) is the principal's participation constraint. It is written only as a function of c_1 and q_1 because symmetry implies that as a function of output the principal gives the same compensation to both agents and receives the same level of output in expectation from both agents. Similarly, the principal's utility is specified at the level $\bar{W}/2$, since he receives half of his utility from agent one.

Equations (33) are agent 1's incentive constraints if he is not switched. In accordance with our assumption that a_1 and b_1 are recommended before the interim stage, the summand is only over consumption, c_1, and outputs, q_1, q_2. The likelihood ratio on the right-hand side of the equation indicates that deviations by the agent only affect output on plot 1.

Equations (34) are agent 1's incentive constraints if he is switched. The likelihood ratios on the right-hand side of the constraint contain two substitutions which might not be readily apparent. When switched, the production function of plot i is $p(q_i|a_i + b_j)$. However, because of symmetry and the assumption that both agents receive the same recommended sequence of actions, we know that $a_1 = a_2$ and $b_1 = b_2$. Substituting, gives the form the ratios take in (34). They indicate that for agent one an interim-stage deviation affects the probability of q_1 and a final-stage deviation affect the probability of q_2. Finally, the remaining equations (35) and (36) are the technology and probability measure constraints, respectively.

Now consider a technological specification where interim and final stage efforts are perfect substitutes in production. That is, $p(q_i|a_i + b_i)$ is the production function for plot i if there is no switch and $p(q_i|a_i + b_j)$ is it if there is a switch. Also, assume that this probability function satisfies the monotone likelihood ratio property (MLRP) and convexity of the distribution function (CDFC). As discussed in section 2.1, these assumptions are sufficient to use the first-order approach in single agent problems, which for our purposes corresponds to checking adjacent constraints. They let us prove the following theorem.

Theorem 5 *When interim and final stage efforts are perfect substitutes in production, and production satisfies MLRP and CDFC, the optimal no-switching symmetric equilibria weakly dominates all switching symmetric equilibria.*

Proof: First, consider the following simplification of the no-switch incentive constraints. Let

$e_i = a_i + b_i$ be the total effort worked by agent i and let

$$\hat{V}(e_i) = \min_{\{a_i, b_i | a_i + b_i = e_i\}} V(a_i) + V(b_i). \tag{37}$$

The solution to the problem is the optimal way for agent i to allocate e_i units of effort across the interim and final stages. The concave function $\hat{V}(e_i)$ is the indirect utility from working e_i units of effort.

Since interim and final stage efforts are perfect substitutes in production any allocation a_i, b_i must respect agent i's ability to reallocate effort between interim and final stage efforts according to equation (37). Consequently, we can frame the incentive constraint (33) solely in terms of e_i. Now recall that the production technology on each plot satisfies MLRP and CDFC, which means that only downward adjacent incentive constraints bind. That is, for a given level of total effort e_i^h, we only need to check the next lowest total effort is e_i^{h-1}. In terms of a_i, and b_i lowering e_i by one unit means either a_i or b_i will be lowered by one unit as determined by $\hat{V}(e_i^h)$ and $\hat{V}(e_i^{h-1})$. (Recall that a_i and b_i are equally spaced on their grids),

Without loss of generality assume that a_i is lowered. Letting h_a and h_b index the recommended a_i and b_i actions respectively, incentive constraints for agent one are

$$\begin{aligned}
&\sum_{c_1, q_1, q_2} \pi(c_1, q_1, q_2, b_1, a_1 | d{=}0)[U(c_1) - V(a_1^{h_a}) - V(b_1^{h_b})] \\
&\quad \geq \sum_{c_1, q_1, q_2} \pi(c_1, q_1, q_2, b_1, a_1 | d{=}0) \frac{p(q_1 | a_1^{h_a - 1} + b_1^{h_b})}{p(q_1 | a_1^{h_a} + b_1^{h_b})} [U(c_1) - V(a_1^{h_a - 1}) - V(b_1^{h_b})].
\end{aligned} \tag{38}$$

If this same agent is switched then he faces the option of deviating on plot one or plot two. Considering only downward adjacent constraints, lowering a_i is identical to equation (38). However, he must also be prevented from deviating on plot two. The incentive constraint on plot two is

$$\begin{aligned}
&\sum_{c_1, q_1, q_2} \pi(c_1, q_1, q_2, b_1, a_1 | d{=}0)[U(c_1) - V(a_1^{h_a}) - V(b_1^{h_b})] \\
&\quad \geq \sum_{c_1, q_1, q_2} \pi(c_1, q_1, q_2, b_1, a_1 | d{=}0) \frac{p(q_1 | a_1^{h_a} + b_1^{h_b})}{p(q_1 | a_1^{h_a} + b_1^{h_b - 1})} [U(c_1) - V(a_1^{h_a}) - V(b_1^{h_b - 1})].
\end{aligned} \tag{39}$$

Any incentive-compatible switching allocation must satisfy constraints (38), (39), plus possibly incentive constraints for non-adjacent actions. No-switching allocations must only satisfy (38). Therefore, no-switching weakly dominates switching. **Q.E.D.**

Long-term arrangements – no-switch allocations – are powerful in this model because they allow incentives to be focused on one plot. In contrast, short-term arrangements – switch allocations – require incentives to induce effort on *both* plots. Holding total plot efforts fixed, a switched agent faces the same tradeoff between consumption and effort on plot 1 as a no-switch agent who considers lowering a_1. However, a switched agent also faces that same tradeoff for b_1 but now on plot 2. The key difference is that a no-switched agent only affects the probability distribution of *one* output but a switched agent affects the distribution of *both* outputs at the margin. The argument is analogous to the analysis of the relative performance with substitutable labor model in Theorem 1. In that model the decision was whether or not to work the agents together on both plots or separately on their own plots. In this model by switching the agents the principal decides how many plots each agent will work.

As in the analysis of section 2.1, the relative performance model, there are cases where it is desirable for agents to work together, that is, be switched in the language of this section. Consider a technology which is of the form $p(q|a,b)$, where a and b are plot-specific interim and final stage efforts, irrespective of the agent. Let the inputs be complementary in the sense that altering one input without altering the other dramatically lowers output, but if both a and b are jointly altered expected output is only moderately lowered.

To be more specific let there be two types of each effort indexed by l and h, and two outputs indexed again by l and h. Assume that l corresponds to low and h corresponds to high, so that $q_l < q_h$, $a_l < a_h$, and $b_l < b_h$. High outputs are more desirable than low outputs and high efforts are less desirable than low efforts. Table 6 lists a technology $p(q|a,b)$ in which inputs are strong complements. The left block of cells describes the probability distribution $p(q|a_l,b)$ while the right block describes $p(q|a_h,b)$. An example of this production function could be an economy where

	a_l			a_h	
	q_l	q_h		q_l	q_h
b_l	.70	.30	b_l	.99	.01
b_h	.99	.01	b_h	.40	.60

Table 6: A production technology, $p(q|a,b)$, which generates switching.

the inputs are types of fertilizer and pesticides which may have strong interactions.

We have not computed an economy like the one above but we conjecture that if the desired effort pair on each plot is a_h, b_h, then switching would be optimal. Consider the following argument. If an agent is not switched then incentives must be sufficiently strong to prevent him from deciding to take the a_l, b_l action. The planner must also prevent him from taking the a_h, b_l and a_l, b_h actions but the extremely high probability of the low output makes preventing these deviations relatively costless.

In contrast, if agents are switched then the agent can not choose the b_l, a_l deviation because he takes the other player's action as given. Then, if either was to deviate on either plot the efforts would be a_l, b_h or a_h, b_l. Since these deviations produce the low output with virtual certainty, they are relatively costless to prevent.

The analysis is analogous to the assessment of the relative performance and group regimes in the model with correlated plot returns. As we saw in section 2.1, relative performance was powerful because deviations in effort were revealed by comparisons of plot outputs. Groups were not desirable in that model because the agents could, by jointly lowering their actions, remove the informativeness of output comparison. Of course, the analogy is not exact; in this section there is no value to comparing outputs. However, in both examples coordination of efforts allowed likelihood ratios to be altered in ways which effort deviations.

So far, the emphasis has been on the tradeoff between coordination and being responsible for one's efforts, a theme emphasized earlier when comparing relative performance with group models. The tradeoff is important but there are additional factors, like information and redistribution of resources which require a different version of the prototype to address. The next subsection demonstrates that switching can improve the information structure of the economy by removing some incentive constraints, and weakening others. It also shows that switching can improve resource reallocation. These results are found in an environment which allows switching to be random and which does not restrict allocations to be symmetric.

4.2 The Role of Information

Consider the following variation on the previous model. First, assume there is a continuum of agents each of whom starts out on a plot of land. For the moment assume that all agents are ex ante identical. In the interim stage replace effort, a, with an interim plot-specific shock, θ. The shock, θ, to a plot is exogenous, drawn from the distribution $h(\theta)$, and only observed by the agent who starts on that plot. Upon observing the state of their plot, agents send a report to the planner. As we will be using the revelation principle, this report is simply on the value of θ.

After receiving the reports, the planner assigns the agents to plots of land which may or may not be the one the agent started out on. Once on their possibly new plot, agents work a final-stage effort, b, which along with the intermediate state of the newly assigned plot, θ, determine the probability distribution of the publicly-observed plot output, q. The probability distribution of output is described by $p(q|b, \theta)$. As before, the interim shock θ and the second-stage efforts, b, are only observed by the agent who was on the plot during that stage. This means that if an agent stays on his original plot he knows the value of θ, but if he is reassigned he does not know the θ on his new plot.

Rather than writing out the program with switching we make some observations which allows us to directly analyze the model. First, consider the following no-switching allocation. If an agent stays on his plot with probability one, then the underlying commodity space would be $\pi(c, q, b|\theta)$, that is the probability distribution of consumption c, output q, and action b, given θ was reported. By the revelation principle the allocation would need to satisfy incentive constraints which induce truthful reporting of θ and then given a truthful report, other constraints which ensure that the agent takes the recommended action. The truth-telling constraints are

$$\begin{aligned}\forall\theta,\ &\sum_{c,q,b}\pi(c,q,b|\theta)[U(c)-V(b)]\\ &\geq\sum_{c,q,b}\pi(c,q,b|\theta')\frac{p(q|\delta(b),\theta)}{p(q|b,\theta')}[U(c)-V(\delta(b))],\ \ \forall\theta'\neq\theta,\ \forall\delta:B\rightarrow B.\end{aligned}\tag{40}$$

Constraints (40) ensure that telling the truth, θ and then taking the action b recommended by π is preferable to lying, *i.e.* sending a report $\theta' \neq \theta$, and then taking any strategy δ of responding to recommended actions b. For more details on this constraint see Myerson (1982) for the original treatment or E.S. Prescott (1995) for an exposition in a similar model.

In addition to truth-telling constraints, the revelation principle requires constraints which ensure that an agent who truthfully reports θ takes the recommended action b. The following constraints do this

$$\begin{aligned}\forall\theta,\ &\sum_{c,q}\pi(c,q,b|\theta)[U(c)-V(b)]\\ &\geq\sum_{c,q}\pi(c,q,b|\theta)\frac{p(q|\hat{b},\theta)}{p(q|b,\theta)}[U(c)-V(\hat{b})],\ \ \forall b,\hat{b}.\end{aligned}\tag{41}$$

Now consider an alternative scheme. Upon receiving agents' reports the planner switches each agent with probability one. Assume the reassignment does not depend in anyway on agents' reports and assume the quality of each agent's assigned plot is randomly drawn from the distribution $h(\theta)$. Further, assume that compensation does not depend in anyway upon an agents report and assume that the planner tells each agent the quality of his new plot θ. Like the previous scheme, an allocation is described by the lottery $\pi(c, q, b|\theta)$, but now θ is the type of the *newly* assigned plot.

This switching scheme, while simple, is surprisingly powerful. First, because agents' utilities do not depend on their reports they truthfully report on the interim state, θ, of their original plot. In this scheme, there are no truth-telling constraints! The only incentive constraints are those on agents actions, which are identical to constraints (41), in the no-switch scheme. Under the switching scheme, agents are information monitors who report truthfully because they are indifferent to the θ they observe or report. The arrangement is essentially a moral-hazard economy, with the added twist that there there is a random, publicly observed, shock to the production technology.

Theorem 6 *Switching and telling the agent the state of his newly assigned plot θ, weakly dominates no switching.*

Proof: A feasible allocation under the just described switching regime must satisfy incentive constraints (41). Feasible allocations under the no-switching regime must satisfy incentive constraints (41) and (40). **Q.E.D.**

Interestingly, it is possible for the planner to do even better than the previous switching scheme. As the following scheme demonstrates, switching not only removes truth-telling constraints but it can mitigate the recommended action constraints (41).

To see this, modify the previous scheme by having the planner not tell the agent the state of his newly assigned plot. Again, there is no truth-telling constraints but now the constraints on the recommended action are

$$\sum_{c,q,\theta} \pi(c, q, b|\theta)h(\theta)[U(c) - V(b)]$$
$$\geq \sum_{c,q,\theta} \pi(c, q, b|\theta)\frac{p(q|\hat{b}, \theta)}{p(q|b, \theta)}h(\theta)[U(c) - V(\hat{b})], \quad \forall b, \hat{b}. \tag{42}$$

Constraints (42) are very similar to constraints (41) used by the other two schemes. The difference is that now the agent does not know the quality θ of his newly assigned plot. Consequently, upon receiving recommended action b, he uses $\pi(c, q, b|\theta)$ and $h(\theta)$ to form a posterior of the quality of his new plot. For example, if only agents assigned to θ_1 quality plots were recommended action b_1, then any agent assigned b_1 would know what type of plot he was assigned, as in the second scheme. However, if agents assigned to plots of different θ were also assigned b_1 then they would not know the quality of their plot with certainty.

We can now prove the following theorem.

Theorem 7 *A switching allocation where the planner does not tell the agent the state, θ, of his new plot weakly dominates a switching allocation where the planner tells the agent the value of θ on his newly assigned plot.*

Proof: The θ in the summand of constraints (42) reflects the inference that agents are making about the quality of their plots based on their recommended action b. Mathematically, it means that each constraint (42) is the weighted sum of several (41) constraints. Consequently, any allocation satisfying (41) for each θ will satisfy (42), but not necessarily vice versa. **Q.E.D.**

The last proof demonstrates that switching might be beneficial not only because it allows for costless revelation of information but because it also allows for costless scrambling of information. As we will see later in a numerical example, this dominance can hold strictly.

We now present an example which illustrates the beneficial role of scrambling. The example also demonstrates that switching can help improve resource allocation, in the sense of allowing different types of people to work different amounts. To study these issues we modify the above economy by allowing heterogeneity of agent types. Let there be a finite number of different types, t, with each type constituting a positive fraction, $f(t)$, of the population. Pareto weights are described by the function $\lambda(t)$, with $\sum_t \lambda(t) = 1$. All agents of a given type are treated ex ante identically.

We take as a starting point, the observation that the planner can costlessly induce truthful revelation of plot qualities, θ, by the schemes described above. Consequently, we use the same commodity space, but with two slight modifications. First, we index allocations by group types and second, we put the assignment of plot qualities directly into the allocation. This latter modification changes the notation and adds a constraint but without altering the problem. Thus, an allocation for type t is a probability distribution $\pi(c, q, b, \theta|t)$ and the program is

Program 8

$$\max_{\pi()} \sum_{c,q,b,\theta,t} \pi(c, q, b, \theta|t) f(t) \lambda(t)[U(c) - V(b)]$$

s.t.

$$\sum_{c,q,b,\theta,t} \pi(c, q, b, \theta|t) f(t)(q - c) = 0, \tag{43}$$

$$\sum_{c,q,\theta} \pi(c, q, b, \theta|t)[U(c) - V(b)] \geq \sum_{c,q,\theta} \pi(c, q, b, \theta|t) \frac{p(q|\hat{b}, \theta)}{p(q|b, \theta)} [U(c) - V(\hat{b})]$$
$$\forall t \in T, \forall b, \hat{b}, \tag{44}$$

$$\forall \bar{\theta}, \quad \sum_{c,q,b} \pi(c, q, b, \bar{\theta}|t) f(t) = h(\bar{\theta}), \tag{45}$$

$$\forall \bar{t}, \bar{q}, \bar{b}, \bar{\theta}, \quad \sum_{c} \pi(c, \bar{q}, \bar{b}, \bar{\theta}|\bar{t}) = p(q|\bar{b}, \bar{\theta}) \sum_{c,q} \pi(c, q, \bar{b}, \bar{\theta}|\bar{t}), \tag{46}$$

$$\forall t, \quad \sum_{c,q,b,\theta} \pi(c, q, b, \theta|t) = 1, \text{ and } \forall c, q, b, \theta, t, \ \pi(c, q, b, \theta) \geq 0. \tag{47}$$

Equation (43) is the resource constraint for the economy. It allows for transfers across types. Equations (44), the incentive constraints, ensure that each agent takes his recommended action. As was discussed earlier, since agents do not know the quality of their newly assigned plots, they form a posterior of plot quality based on their type and recommended action. If others of their same type are assigned to a different type of plot but receive the same recommended action then the agent forms a non-trivial posterior over what type of plot he is working. The summation over θ incorporates these expectations. Equation (45) ensures that the agents are assigned to plot types in proportion to the distribution of plot types. The rest of the constraints are standard to the linear programming approach and have been described in similar models earlier in the paper.

The following example demonstrates the scrambling and resource reallocation benefits of switching.

Numerical Example 5

Let there are two types, t_1 and t_2, who comprise 80% and 20% of the population respectively. The Pareto weights on the types are, $\lambda(1) = .2$, and $\lambda(2) = .8$. In this economy there is a small fraction of the population who receives proportionally much more weight than the rest of the population.

Preferences are separable in consumption and effort. The consumption grid is of length .02 over the range 0 to 2. Utility from consumption is $U(c) = c^{.5}/.5$. Agents may only choose one of three efforts, b_1, b_2, or b_3. The effort portion of the utility function is described by $V(b_1) = V(b_2) = -2$, and $V(b_3) = -\sqrt{2}$. Recalling that preferences are $-V(b)$, efforts b_1 and b_2 give equal utility but are different actions. Effort b_3 gives the least utility and corresponds to hard or unpleasant work.

There are three different types of plots, indexed by θ_1, θ_2, and θ_3. Plot types are random and drawn from the distribution $h(\theta_1) = h(\theta_2) = .4$, and $h(\theta_3) = .2$. Each type of plot may produce either a low output, $q_l = 0$ or a high output, $q_h = 1$. The probability distribution of the outputs is a function of the plot type and the action and is described by $p(q|b,\theta)$. Output on each plot is independent of other plots. Table 7 describes the $p(q|b,\theta)$ function used in the example.

	θ_1			θ_2			θ_3	
	q_l	q_h		q_l	q_h		q_l	q_h
b_1	.99	.01	b_1	.70	.30	b_1	.99	.01
b_2	.70	.30	b_2	.99	.01	b_2	.99	.01
b_3	.40	.60	b_3	.40	.60	b_3	.99	.01

Table 7: A production technology, $p(q|b,\theta)$, which generates switching.

The first two technologies, indexed by θ_1 and θ_2, are the most productive as long as b_3 is worked. The two types of plots are identical except that $b_i, i = 1, 2$ has a different effect on each plot. If b_i is worked on a $\theta_i, i = 1, 2$, plot then the plot is extremely unproductive. The third plot is so unproductive that efforts do not have any effect at all.

Table 8 lists the solution to the example. The top half of the table lists the optimal allocation for type 1s and the bottom half lists it for type 2s. The second column contains the probability of being assigned to each land type and the third column contains the action taken conditioned on the land assignment. Finally, fourth and fifth columns list expected consumption given output. Expected consumption is listed because the solution contained lotteries over adjacent consumption

	Type 1's allocation			
	$\pi(\theta_i)$	Action	$E(c\|q=0)$	$E(c\|q=1)$
θ_1	0.5	b_3	0	.43
θ_2	0.5	b_3	0	.43
θ_3	0	-	-	-

	Type 2's allocation			
	$\pi(\theta_i)$	Action	$E(c\|q=0)$	$E(c\|q=1)$
θ_1	0	-	-	-
θ_2	0	-	-	-
θ_3	1	b_1 or b_2	1.37	1.36

Table 8: Optimal Contract as a Function of the Type and Land Assignment

grid points.

Type 1 agents are assigned randomly to the productive plots, θ_1 and θ_2, while the type 2 agents are always assigned to the low quality plots, θ_3. On the productive plots the type 1 agents are being induced to work hard, b_3, by a contract that gives them 0 consumption if output is low and .43 units of consumption if output is high. Type 2 agents work either b_1 or b_2, since both are just as unproductive and gives the same utility. They receive large consumption transfers and are also fully insured over outputs because there is no incentive problem on their efforts.[8] Type 2's are truly the idle rich.

Scrambling over θ_i, $i = 1, 2$, of type 1's land assignment mitigates incentives. Fifty percent of the time this type of agent is assigned to a θ_1 plot and the other half of the time he is assigned to a θ_2 plot. Since he does not know the type of his assigned plot he forms expectations based on what he knows. His knowledge at that point is his type and his recommended action. He infers from the equilibrium distribution that since he is a type 1 and was recommended action b_3 there is a 50% chance he has been assigned to a θ_1 plot and a 50% chance he has been assigned to a θ_2 plot. He knows for sure that he is not working a θ_3 plot because only type 2s are assigned to θ_3 plots. Now consider the choice facing this agent after being recommended action b_3. If he takes either action b_1 or b_2 the $V(a)$ portion of his utility increases but there is an 85% chance of receiving the low output, and the dependence of consumption on output is sufficient to preclude this option.

[8]The difference listed in the table between high output consumption and low output consumption exists only because of the discreteness of the consumption grid.

The 85% chance of the low output is directly due to the scrambling of land assignments. Consider the case where the planner tells a type 1 agents the state of his plot. If he is assigned to a θ_1 plot and is recommended action b_3 then an incentive compatible contract must stop him from taking action b_1 or b_2. The very unproductive action, b_1, is easy to prevent but the other action, b_2, is much harder to detect since it produces the low output 70% of the time. Compare this number with the 85% chance of the low output if the agent deviates under the scrambling regime. It is lower so consequently, the incentives necessary to make b_3 incentive compatible are costlier than when there is scrambling. This result can be confirmed if the above program is solved when there is no θ_2 plots, and instead 80% of the plots are θ_1 while the rest still are θ_3. The optimal allocation in this economy gives up on implementing the b_3 action because it is too costly to make that action incentive compatible.

In the examples, there is another way in which switching is beneficial in the example. Without switching some type 2s would work high quality plots while some type 1s would work low quality plots. Such an assignment is clearly not optimal because given the distribution of planner's weights, it is desirable for type 1s to work hard and type 2s to not work hard. Switching ensures hard workers are assigned to plots where their hard work is productive and idle workers are assigned to plots where the loss from their lack of work is relatively small.

Let us add some final words to our discussion of long-term versus short-term arrangements. Switching was shown to be powerful because of the ability to costlessly elicit information. Once the information was obtained by the planner, switching also allowed for scrambling of assignments and efficient reallocation of inputs. However, if the first model was combined with this model, possibly by letting an initial action, a, randomly determine θ, then information might not be costless to elicit. If there was moral-hazard on the initial action then there would need to be some dependence of utility on the report or agents would not work high levels of a. Switching behavior should be considerably more complicated in this model. Other alternative directions, include modifying the models to make reporting from the agents to the planner costly or limited. Little work has been done on models with costly and limited communication, though E.S. Prescott (1995) is a start.

5 Conclusion

Our goal in this paper is to study the fundamental economic question posed in Coase (1937): Why are there firms? Motivated by observations of firms in several economies, prototypes were developed to study three characteristics of firms. We first studied the joint decision making power of firms. Firms were treated as collections of individuals who could make and enforce their own rules within the firm. These group regime models were compared with relative performance models. A substantial contribution to the relative performance literature was made by demonstrating that the distribution of wealth was an important factor in determing optimal organizational structure.

Next, the group model was extended to a general equilibrium setting. It was shown how, with an appropriate choice of the commodity space and consumption set, standard results from general equilibrium could be used. Finally, it was demonstrated that under certain conditions a separate smaller market could be opened within each group and operate simultaneously to the group's trading in across-group market. The section demonstrates how decentralized mechanisms may coexist at different levels of the economy.

The remaining sections of the paper used an assignment definition of a firm to analyze two other common characteristics of firms. First, an information monitoring model was developed where being assigned to the same plot of land as another agent allowed for the revelation of information. Numerical examples were provided which demonstrated the role that the distribution of wealth plays in determing the optimal organizational structure. Economies were also provided in which optimal land assignments resembled the land usage pattern of the cropping groups in Aurepalle.

The second assignment characteristic of firms studied was their long-term nature. In multi-stage models, the problem was framed in terms of interim reassignments of labor. The idea is that agents who are not reassigned are in a firm while reassigned agents are participating, in a sense, in a spot market. Examples were provided which demonstrated that short-term arrangements allowed for the planner to obtain information and use it judiciously when reassigning labor. In another class of production functions, long-term arrangements were shown to be beneficial through their potential to make clear which individual was responsible for which output.

A Computational Appendix

This appendix describes the algorithm used to compute solutions to the land assignment problem, Program 6 in Section 3. First, we describe how assumptions on preferences and technology allow the program to be split into subproblems. Second, we describe how to compute the subproblems.

Three important assumptions were made in numerical examples 3 and 4. First, utility is linear in effort, that is, $U(c_i) - k_e e_1$. Second, plot returns are independent. Third, the principal is risk neutral. The latter two assumptions are used to prove the following result which is originally due to Holmström (1982).

Theorem 8 *If plot returns are independent and the principal is risk-neutral then consumption is only a function of the output on an agent's privately worked plot.*

Proof: For simplicity, we consider only the $(1,1,1)$ land assignment, but similar arguments may also be made for other land assignments. Recall that under the $(1,1,1)$ assignment agent one works plot one alone, agent two works plot two alone, and both agents jointly work plot three. Let $\pi(c_1, c_2, q_1, q_2, q_3, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})$ be a solution to Program 6. The allocation may be used to construct the two following probability distributions

$$\pi_1(c_1, q_1, \mathbf{e_{1\bullet}}) = \sum_{c_2, q_2, q_3, \mathbf{e_{2\bullet}}} \pi(c_1, c_2, q_1, q_2, q_3, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}), \tag{48}$$

$$\pi_2(c_2, q_2, \mathbf{e_{2\bullet}}) = \sum_{c_2, q_2, q_3, \mathbf{e_{1\bullet}}} \pi(c_1, c_2, q_1, q_2, q_3, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}), \tag{49}$$

The probability distribution π_i, $i = 1, 2$ is the unconditional probability of $(c_i, q_i, \mathbf{e_{i\bullet}})$.

Agent one's incentive constraint in Program 6 is equation (27) which we repeat for convenience

$$\begin{aligned} \forall \mathbf{e_{1\bullet}}, \quad & \sum_{\mathbf{c}, \mathbf{q}, \mathbf{e_{2\bullet}}} \pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}})(U(c_1) - V(e_{11} + e_{13})) \\ \geq & \sum_{\mathbf{c}, \mathbf{q}, \mathbf{e_{2\bullet}}} \pi(\mathbf{c}, \mathbf{q}, \mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}) \frac{p(\mathbf{q}|\hat{e}_{11}, e_{22}, e_{13} + e_{23})}{p(\mathbf{q}|e_{11}, e_{22}, e_{13} + e_{23})}(U(c_1) - V(\hat{e}_{11} + e_{13})), \quad \forall e_{11}, \hat{e}_{11}, \end{aligned} \tag{50}$$

Because plot returns are independent the likelihood ratio on the right-hand side of equation (50) simplifies to $\frac{p(q_1|\hat{e}_{11})}{p(q_1|e_{11})}$. Consequently, coefficients on π in (50) are only a function of c_1, q_1, and $\mathbf{e_{1\bullet}}$ so (48) may be used to rewrite the incentive constraints as

$$\sum_{c_1, q_1} \pi_1(c_1, q_1, \mathbf{e_{1\bullet}})(U(c_1) - V(\mathbf{e_{1\bullet}})) \geq \sum_{c_1, q_1} \pi_1(c_1, q_1, \mathbf{e_{1\bullet}}) \frac{p(q_1|\hat{e}_{11})}{p(q_1|e_{11})}(U(c_1) - V(\hat{e}_{11} + e_{13})), \ \forall e_{11}, \hat{e}_{11}. \tag{51}$$

Equation (51) implies that any randomization in c_1 obtained through realizations of q_2 and q_3 can be obtained through dependence on q_1, without altering agent one's incentive constraints. The same argument can be made for agent two.

The only other way q_2 and q_3 may affect c_1 is through the resource constraint. However, the principal is risk-neutral so he can costlessly insure against any fluctuations in aggregate output. Consequently, consumption, c_i, of agent i need only be a function of q_i (and possibly $\mathbf{e_{i\bullet}}$). **Q.E.D.**

The next simplification uses the assumption that disutility is linear in effort. Consider the incentive constraint (51). Theorem 8 showed that q_2, q_3, and $\mathbf{e_{2\bullet}}$ are not relevant for determining c_1. It can be shown that effort on the group plot, e_{13} is not relevant either. To see this, observe that the only portion of the incentive constraint affected by e_{13} is the level of utility; it does not affect the likelihood ratio. However, it's effect on the level of utility does not have a real impact on the incentive constraint because e_{13} cancels out on both sides of the constraint. Efforts on the group plot do not affect an agent's incentive to work on his own plot.

This result along with Theorem 8 imply that any land assignment may be broken into separate components: plots worked privately by agent one, plots worked privately by agent two, and plots worked by the group. The only connection between the components will be through each component's contribution to agents' utilities and to the principal's utility. Efforts on group plots will determine the amount of resources that the principal may distribute to the agents but they will not affect agent's incentive to work his own plot nor will they allow the principal to make a better inference on agent's private efforts.

The simplifications are helpful but the subproblems themselves may still be large. In particular, for a $(3,0,0)$ land assignment, where agent one works all three plots by himself, the number of feasible action combinations can make this subproblem very large. Solving this subproblem globally, that is by one large linear program, is not feasible for the dimensions necessary to make the land assignment problem interesting. Consequently, we make use of the dynamic programming technique developed in Prescott (1995) to compute a more complicated class of private information models. The technique will use the previous observations in ways which will soon be made evident.

We break the problem into two stages. As usual with dynamic programming methods we work backwards. In the second stage we calculate the least expensive way for the principal to implement a given effort while delivering a certain level of utility to the agent. For each effort level we make this calculation for all feasible utilities which may accrue to the agent at that level of effort; essentially, we are calculating Pareto frontiers for each level of effort. Then, in the first stage, the optimal combination of second stage Pareto frontiers is chosen.

For simplicity, we only present the second stage problem for the case of agent one working one plot privately. As before, his effort on his private plot is denoted e_{11}. Let v_1 be the utility accruing to agent one *from his privately worked plot*, and let $w_1(v_1, e_{11})$ be the utility accruing to the principal *also from agent one's privately worked plot* given e_{11} is recommended and v_1 is promised to the agent. The choice variable is $\pi(c_1, q_1|e_{11})$, the joint probability of consumption c_1 and output q_1, given effort e_{11} is to be implemented. The function $w_1(v_1, e_{11})$ is obtained by solving the following problem for each v_1 and e_{11}.

Stage Two Program

$$\max_{\pi(\cdot)} \sum_{c_1,q_1} \pi(c_1, q_1|e_{11})(q_1 - c_1)$$

$$\text{s.t.} \quad \sum_{c_1,q_1} \pi(c_1, q_1|e_{11})(U(c_1) - k_e e_{11}) \geq v_1, \tag{52}$$

$$\sum_{c_1,q_1} \pi(c_1, q_1|e_{11})(U(c_1) - k_e e_{11}) \geq \sum_{c_1,q_1} \pi(c_1, q_1|e_{11}) \frac{p(q_1|\hat{e}_{11})}{p(q_1|e_{11})}(U(c_1) - k_e \hat{e}_{11}), \quad \forall \hat{e}_{11}, \tag{53}$$

$$\forall \bar{q}_1, \ \sum_{c_1} \pi(c_1, \bar{q}_1|e_{11}) = p(\bar{q}_1|e_{11}), \tag{54}$$

$$\sum_{c_1,q_1} \pi(c_1, q_1|e_{11}) = 1, \quad \text{and} \ \forall c_1, q_1, \ \pi(c_1, q_1|e_{11}) > 0. \tag{55}$$

The program is similar in structure to the earlier programs. Equation (52) is the promised utility constraint. It guarantees that agent one receives v_1 utils from consumption and from effort on plot one. Equations (53) are the incentive constraints. They ensure that e_{11} is incentive compatible. Constraints (54) ensure that the choice variable is consistent with the technology, and equation (55) is the probability measure constraint. Stage two programs for agent two are similar as are the programs for cases with more than one privately worked plot.

The computational advantage and disadvantage of the dynamic programming approach should be evident at this point. By fixing effort, the number of variables and constraints in a stage two program is small so the program computes quickly. The disadvantage is that a large number of such programs must be computed, one for each (v_1, e_{11}). The number of these will depend on the size of the effort, e_{11}, and promised utility, v_1, grids.

The final subproblem to compute is the one corresponding to the group plot. Let w_g and v_g denote the principal's and each agent's utility from the group plot. As shown in Theorem 8 agents do not receive their consumption from the group plot. Consequently, the calculation does not require solving an optimization problem but instead calculating the following functions

$$v_g = -k_e e_{13} = -k_e e_{23},$$
$$w_g = \sum_{q_3} p(q|e_{13} + e_{23})q.$$

The two agents are identical and work equal amounts on the group plot so they both receive the same disutility, v_g, from the group plot.

The stage two calculations create value functions, $w_1(\cdot)$, $w_2(\cdot)$, and $w_g(\cdot)$ which depend on promised utilities and recommended efforts. For computational purposes it is helpful to rewrite the value functions as vectors. We do this by writing $(w_1, v_1, e_{11}) = (w_1(v_1, e_{11}), v_1, e_{11})$.

The stage one program chooses the optimal probability distribution over promised utilities v_1, v_2, v_g and recommended efforts $\mathbf{e_{1\bullet}}, \mathbf{e_{2\bullet}}$ which is consistent with the land assignment. If the land assignment is $(1,1,1)$, the choice variables are probabilities, $\pi_1(w_1, v_1, e_{11})$, $\pi_2(w_2, v_2, e_{22})$, and $\pi_g(w_g, v_g, e_{13})$. (Recall that $e_{13} = e_{23}$ so we can drop reference to e_{23} in v_g.) If, instead, the land assignment is $(2,0,1)$, then the choice variables would be $\pi_1(w_1, v_1, e_{11}, e_{12})$, and $\pi_g(w_g, v_g, e_{13})$. Since agent two does not privately work any plots in the $(2,0,1)$ case the π_2 component would be dropped.

For simplicity we only describe the stage one program for the $(1,1,1)$ land assignment. Letting λ_1 and λ_2 be the Pareto weights for agents one and two, respectively, the stage one program is

Stage One Program

$$\max_{\pi_1(\cdot),\pi_2(\cdot),\pi_g(\cdot)} \sum_{w_1,v_1,e_{11}} \pi_1(w_1, v_1, e_{11})\lambda_1 v_1 + \sum_{w_g,v_g,e_{13}} \pi_g(w_g, v_g, e_{13})(\lambda_1+\lambda_2)v_g + \sum_{w_2,v_2,e_{22}} \pi_2(w_2, v_2, e_{22})\lambda_2 v_2$$

$$\text{s.t.} \quad \sum_{w_1,v_1,e_{11}} \pi_1(w_1, v_1, e_{11})w_1 + \sum_{w_g,v_g,e_{13}} \pi_g(w_g, v_g, e_{13})w_g + \sum_{w_2,v_2,e_{22}} \pi_2(w_2, v_2, e_{22})w_2 \geq \bar{W}, \tag{56}$$

$$\sum_{w_1,v_1,e_{11}} \pi_1(w_1, v_1, e_{11}) = 1,$$

$$\sum_{w_g, v_g, e_{13}} \pi_g(w_g, v_g, e_{13}) = 1,$$
$$\sum_{w_2, v_2, e_{22}} \pi_2(w_2, v_2, e_{22}) = 1, \tag{57}$$

and conditions that the choice variable are all non-negative.

Equation (56) guarantees that the principal receives $\bar{W}$ utils. It is the only constraint which connects the subproblems. Equations (57) are the probability measure constraints. The program has a large number of variables for fine grids of v_1 and v_2, but there are only four constraints, which makes it relatively easy to compute.

To compute numerical examples 3 and 4 the following steps were taken. Stage two calculations were done for agent one working one, two, and three plots alone. Since agents' preferences are identical the same calculations were used to create the value function for agent two's subproblem. Next, the value function on the group plot, $w_g(\cdot)$, was calculated for one, two, and three plots.

After the stage two calculations were completed, the stage one program was solved for each land division. As described earlier the land division determines which combinations of stage two problems are feasible. For example, if the land division was $(2, 0, 1)$ then only values corresponding to the two-plots-worked-alone case and the one-group-plot case were feasible for the stage one problem. Comparison of the values of the objective function across land assignments determines the optimal one. Finally, it is necessary to check that $\sum_j e_{ij} \leq 12$, that is, no agent works more than his total effort. This latter constraint is not guaranteed to hold in the stage one computation. However, if the global solution satisfies this constraint then it is an optimum. In the numerical examples, parameters were purposely chosen so that this constraint was not violated.

References

[1] Benjamin, Dwayne. "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models." *Econometrica*, 60 (1992): 287-322.

[2] Carter, Michael R. "Risk Sharing and Incentives the Decollectivization of Agriculture." *Oxford Economic Papers*, 39 (1987): 577-95.

[3] Chiappori, Pierre-André. "Collective Labor Supply and Welfare." *Journal of Political Economy*, 100 (1992): 437-67.

[4] Coase, R H. "The Nature of the Firm." *Economica*, New Series, 4 (1937): 386-405.

[5] Cochrane, John H. "A Simple Test for Consumption Insurance." *Journal of Political Economy*, 99 (1991): 957-76.

[6] Deaton, Angus. "On Risk, Insurance, and Intra-village Smoothing." Manuscript, 1993.

[7] Debreu, Gerard. "Valuation Equilibrium an Pareto Optimum." *Proceedings of the National Academy of Sciences* 40 (1954): 588-92.

[8] Feder, Gershon, et al. "Credit's Effect on Production in Chinese Agriculture." World Bank, 1991.

[9] Gaynor, Martin and Paul Gertler. "Moral Hazard and Risk Spreading in Partnerships." *RAND Journal of Economics* 26 (1995): 591-613.

[10] Gorman, W.M. "Community Preference Fields." *Econometrica* 21 (1953): 373-90.

[11] Grossman, Sanford and Oliver Hart. "An Analysis of the Principal Agent Problem." *Econometrica* 51 (1983): 7-45.

[12] Hammond, Peter J. "Markets as Constraints: Multilateral Incentive Compatibility in Continuum Economies." *Review of Economic Studies* 54 (1987): 399-412.

[13] Harris, Milton, and Robert Townsend. "Resource Allocation under Asymmetric Information." *Econometrica* 49 (1981): 33-69.

[14] Hart, Oliver and Bengt Holmström. "The Theory of Contracts." In *Advances in Economic Theory: Fifth World Congress*. (Truman F. Bewley Ed.) Cambridge, England: Cambridge University Press, 1987: 71-155.

[15] Haubrich, Joseph G. "Optimal Financial Structure in Exchange Economies." *International Economic Review* 29 (1988): 217-35.

[16] Holmström, Bengt. " Moral Hazard in Teams." *Bell Journal of Economics* (1982): 324-40.

[17] Holmström, Bengt and Paul Milgrom. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica* 55 (1987): 303-28.

[18] Holmström, Bengt and Paul Milgrom. "Regulating Trade Among Agents." *Journal of Institutional and Theoretical Economics* 146 (1990): 85-105.

[19] Hornstein, Andreas and Edward C. Prescott. "Insurance Contracts as Commodities: A Note." *Review of Economic Studies* 58 (1991): 917-28.

[20] Itoh, Hideshi. "Incentives to Help in Multi-agent Situations." *Econometrica* 59 (1991): 611-36.

[21] Itoh, Hideshi. "Coalitions, Incentives, and Risk Sharing." *Journal of Economic Theory* 60 (1993): 410-27.

[22] Lang, Kevin and Peter-John Gordon. "Partnerships as Insurance Devices: Theory and Evidence." *RAND Journal of Economics* 26 (1995): 614-29.

[23] Mace, Barbara J. "Full Insurance in the Presence of Aggregate Uncertainty." *Journal of Political Economy*, 99 (1991): 928-56.

[24] Meyer, Carrie A. "Agrarian Reform in the Dominican Republic: An Associative Solution to the Collective/Individual Dilemma." *World Development*, 17 (1989): 1255-67.

[25] Mueller, Rolf A.E. and Robert M. Townsend. "Cropping Groups as Firms." Manuscript 1994.

[26] Mueller, Rolf A.E. and Edward Simpson Prescott. "What are the Bullocks Doing Today? Towards an Institutional Analysis of the Transactions in Village Markets for Bullock Services in the Semi-arid Tropics of India." Manuscript, 1996.

[27] Myerson, Roger B. "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems." *Journal of Mathematical Economics* 10 (1982): 67-81.

[28] Prescott, Edward C. and Robert M. Townsend. "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard." *Econometrica* 52 (1984a): 21-45.

[29] Prescott, Edward C. and Robert M. Townsend. "General Competitive Analysis in an Economy with Private Information." *International Economic Review* 25 (1984b): 1-20.

[30] Prescott, Edward Simpson. "Communication in Models with Private Information: Theory, Applications, and Computation." Ph.D. Dissertation, University of Chicago, August 1995.

[31] Ramakrishnan, R.T.S. and A.V. Thakor. "Cooperation versus Competition in Agency." *Journal of Law, Economics, & Organization* 7 (1991): 248-83.

[32] Stokey, Nancy L., Robert E. Lucas, with Edward C. Prescott. *Recursive Methods in Economic Dynamics.* Cambridge, MA: Harvard University Press, 1989.

[33] Tirole, Jean. "Collusion and the Theory of Organizations." In *Advances in Economic Theory Sixth World Congress, Vol. II* (Jean-Jacques Laffont Ed.), Cambridge, England: Cambridge University Press, 1992.

[34] Townsend, Robert M. *The Medieval Village Economy: A Study of the Pareto Mapping in General Equilibrium Models.* Princeton, N.J.: Princeton University Press, 1993.

[35] Townsend, Robert M. "Risk and Insurance in Village India." *Econometrica*, 62 (1994): 539-91.

[36] Udry, Christopher. "Gender, Agricultural Production and the Theory of the Household." Manuscript, 1994.