Ec717a: Robustness and Linear Contracts (Carroll AER2015)

Dilip Mookherjee

Boston University

Ec 717a, 2020: Lecture 3

Introduction

- Recall the simple contract design problem for a single risk averse agent with moral hazard (Holmstrom 1979, Grossman-Hart 1983)
- Optimal contract is quite complicated, and rarely linear

Introduction

- Recall the simple contract design problem for a single risk averse agent with moral hazard (Holmstrom 1979, Grossman-Hart 1983)
- Optimal contract is quite complicated, and rarely linear
- When risk-sharing is unnecessary (both P and A are risk-neutral) but there is limited liability (besides moral hazard) – Innes (JET 1990) shows the optimal contract is piece-wise linear $w(y) = \min\{\alpha y, \bar{y}\}$

Introduction

- Recall the simple contract design problem for a single risk averse agent with moral hazard (Holmstrom 1979, Grossman-Hart 1983)
- Optimal contract is quite complicated, and rarely linear
- When risk-sharing is unnecessary (both P and A are risk-neutral) but there is limited liability (besides moral hazard) – Innes (JET 1990) shows the optimal contract is piece-wise linear $w(y) = \min\{\alpha y, \bar{y}\}$
- However real-world contracts in moral hazard settings are often linear (e.g., sharecropping, uber drivers, authors, executives)
- The share of the agent does not vary finely with detailed features of the environment such as the production function, agent disutility or nature of uncertainty (e.g., 'standard' contracts)

Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture

By H. Peyton Young and Mary A. Burke*

Survey data suggest that cropsharing contracts exhibit a much higher degree of uniformity than is warranted by economic fundamentals. We propose a dynamic model of contract choice to explain this phenomenon. Landowners and tenants recontract periodically, taking into account expected returns as well as conformity with local practice. The resulting stochastic dynamical system is studied using techniques from statistical mechanics. The most likely states consist of patches *where contractual terms are nearly uniform, separated by* boundaries *where the terms shift abruptly. These and other predictions of the model are borne out by survey data on agricultural contracts in Illinois.* (*JEL* J43,C73)

> [T]he constraining force of custom and public opinion ... resembled the force which holds rain-drops on the lower edges of a window frame: the repose is complete till the window is violently shaken, and then they fall together.
>
> —Alfred Marshall

Economists have long been puzzled by the extent to which local custom, rather than competition, shapes the terms of certain kinds of contracts. A well-known example is cropsharing contracts, whereby a landlord leases his farm to a tenant laborer in return for a fixed share of the crops. The high degree of uniformity in the terms of such contracts has attracted the attention (though by no means the approval) of almost all writers on the subject, both ancient and modern. In speaking of the system then prevalent in Italy and France, for example, John Stuart Mill remarks:

> This proportion ... is usually (as is implied by the words *metayer, mezzaiuolo,* and *medietarius*) one-half. There are places, however, such as the rich volcanic soil of the province of Naples, where the landlord takes two-thirds ... . Whether the proportion is two-thirds or one-half, it is a fixed proportion, not variable from farm to farm, or from tenant to tenant (Mill, 1848 p. 303).

Similarly, in a more recent study of contract forms in West Bengal, Ashok Rudra writes:

> the proportion 50:50 for paddy shows a great resilience in that it is known to have existed for a long time in this state ... irrespective of soil conditions, improved or backward methods of cultivation, and other factors which could be expected to affect the profitability of farm business (Rudra, 1975 p. A58).

For those who find these observations disturbing or implausible, there are two comforting explanations. One is that cropsharing is largely a feature of pre-economic, custom-bound cultures; in modern societies, contracts are surely structured more rationally and are governed by competitive forces. Alfred Marshall seems to have been of this view, asserting that the sway of custom is a feature of "primitive times and backward countries."

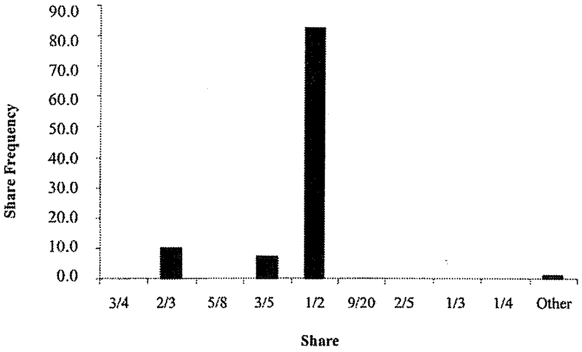

FIGURE 1. CROP SHARE FREQUENCIES IN ILLINOIS: TENANT'S SHARE OF THE CORN CROP (Frequencies in percent)

Source: Illinois Cooperative Agricultural Extension Service *Farm Leasing Survey*, 1995.

Whats Missing in the Model?

- *Effort Arbitrage:* In a risk-sharing-incentives context, Holmstrom-Milgrom (Ecta 1987) model the idea that the agent has a lot of opportunity to 'arbitrage' effort across points of time
- But their model delivers linearity of optimal contracts only when combined with restrictive assumptions (CARA utility, continuous time, Brownian motion)
- *Robustness:* In a setting with risk neutrality and limited liability, Carroll (2015) models a longstanding idea (Hurwicz and Shapiro (BJE 1978)) that P does not know much about the environment, e.g. the exact set of technologies available to A, and may not be able to formulate a Bayesian prior over these
- Could evaluate contracts in terms of their worst-case performance, i.e. minimum payoff guarantees

Non-Bayesian Model

- Real world phenomenon: P often does not know many features of the environment, is unable to form priors and behave like a Bayesian decision-makers
- So P seeks to maximize worst-case-scenario profits
- Similar approach is used in computer science or electrical engineering in selecting algorithms
- Related also to ideas of ambiguity-aversion, and 'satisficing'

Intuition

- 'Simple' (linear) contracts are good because they are robust/reliable
- Trying sophisticated non-linear contracts optimal relative to some prior, could backfire if P makes 'a mistake' in forming the prior
- With linear contracts, expected net returns of P and A move together, so if A does something unexpected to raise her own return, it will also raise P's return — so P is 'protected'

Main Features of Carroll Model

- Risk neutral agent A, subject to limited liability, zero outside option payoff
- Selects costly effort c, which determines probability distribution F over compact set of possible outputs $Y \subseteq \Re$; normalize min Y to 0.
- P offers contract a continuous function $w(y) : Y \rightarrow \Re^+$ (limied liability)
- P has limited information about the 'technology' available to A:
  - knows that A definitely has access to a given set $\mathcal{A}_0$
  - does not know what additional options A has, i.e., A has access to some $\mathcal{A}$ where $\mathcal{A}_0 \subseteq \mathcal{A}$
- *Nontriviality Assumption* There exists $(F, c) \in \mathcal{A}_0$ such that $E_F[y] - c > 0$

Contracts and Payoffs

- Ex post payoff (if y realized): $w(y) - c$ for A, $y - w(y)$ for P
- Timeline: (i) P offers $w(.)$, (ii) A selects $(F, c) \in \mathcal{A}$, (iii) y and payoffs realized
- No participation decision for A incorporated (or assume that $(\delta_0, 0) \in \mathcal{A}_0$, where δ_y denotes degenerate distribution concentrated on y)

Contracts and Payoffs, contd.

- Agent's action choice: (F, c) chosen from $\mathcal{A}$ to maximize $E_F[w(y)] - c$
- Denote resulting payoff of A by

$$V_A(w|\mathcal{A}) = \max_{(F,c)\in\mathcal{A}} E_F[w(y)] - c$$

- P's payoff conditional on $\mathcal{A}$ is

$$V_P(w|\mathcal{A}) = \max_{(F,c)\in A^*(w|\mathcal{A})} E_F[y - w(y)]$$

  where $A^*(w|\mathcal{A})$ denotes the set of optimal choices for A (break ties in P's favor)

- P selects w to maximize worst-case-profit:

$$V_P(w) = \inf_{\mathcal{A}\supseteq\mathcal{A}_0} V_P(w|\mathcal{A})$$

Theorem

There exists linear contract $w(y) = \alpha y$ *for some* $\alpha \in [0, 1]$ *that maximizes* $V_P(.)$.

What P can Guarantee with a Linear Contract (Lemma 2.3)

- Given any $w(.)$, P knows that A will attain an exp payoff of at least $V_A(w|\mathcal{A}_0)$, implying (if (F, c) is actually chosen):

$$E_F[w(y)] \geq (E_F[w(y)] - c \geq) V_A(w|\mathcal{A}_0)$$

- If $w^*(y) = \alpha y$, then $y - w^*(y) = (1 - \alpha)y = \frac{1-\alpha}{\alpha} w^*(y)$
- Hence

$$\begin{aligned} V_P(w^*) &\geq \frac{1-\alpha}{\alpha} V_A(w^*|\mathcal{A}_0) \\ &= \frac{1-\alpha}{\alpha} \max_{(F,c)\in\mathcal{A}_0} \{\alpha E_F[y] - c\} \end{aligned}$$

- Finally observe RHS is exactly P's profit if A's feasible set is $\mathcal{A}_0$, hence

$$V_P(w^*) = \frac{1-\alpha}{\alpha} \max_{(F,c)\in\mathcal{A}_0} \{\alpha E_F[y] - c\}$$

Structure of Rest of Proof

- Show that given any nonlinear contract, there exists a linear contract which generates at least much worst case profit for P
- Use a separating hyperplane argument
- First step is to show that wlog can widen range of options for A to include every probability distribution F over Y

Lemma 2.2 *Let* w *be any eligible contract (satisfying* $V_P(w) > 0$, $V_P(w) \geq V_P(0)$*, where* 0 *is the null contract), which is non-null. Then*

$$V_P(w) = \underline{\pi} \equiv \inf E_F[y - w] \quad s.t. F \in \Delta(Y), E_F(w) \geq V_A(w|\mathcal{A}_0)$$

If the minimum is attained at F^**, then* $E_{F^*}[w] = V_A(w|\mathcal{A}_0)$*.*

Widening the feasible set as much as possible: allowing the agent here to select *any* F, besides the necessary condition $E_F(w) \geq V_A(w|\mathcal{A}_0)$

Since we are taking an upper bound of A's feasible set, obvious that $V_P(w) \geq \underline{\pi}$

So need to show that $V_P(w) \leq \underline{\pi}$, i.e., that P cannot guarantee anything strictly above this.

Steps in Proof of Lemma 2.2

- If P can guarantee something better than $\underline{\pi}$, take the (approximate) solution F^* in the defined problem, and show P's profit could be close to it in some cases
- Suppose A did have access to (only) the option $(F^*, c = 0)$, besides $\mathcal{A}_0$: would she choose it?

Proof of Lemma 2.2, contd.

- If F^* does not put all its weight on y's that maximize w on Y, we can select a distribution F^{**} close to it which assigns a little bit more weight to outputs generating higher wages

Proof of Lemma 2.2, contd.

- If F^* does not put all its weight on y's that maximize w on Y, we can select a distribution F^{**} close to it which assigns a little bit more weight to outputs generating higher wages
- In that case $E_{F^{**}}[w] > E_{F^*}[w] \geq V_A(w|\mathcal{A}_0)$, and A would choose $(F^{**}, c = 0)$ if that is the only option apart from the set $\mathcal{A}_0$, resulting in payoff for P arbitrarily close to $\underline{\pi}$

Proof of Lemma 2.2, contd.

- If F^* does not put all its weight on y's that maximize w on Y, we can select a distribution F^{**} close to it which assigns a little bit more weight to outputs generating higher wages
- In that case $E_{F^{**}}[w] > E_{F^*}[w] \geq V_A(w|\mathcal{A}_0)$, and A would choose $(F^{**}, c = 0)$ if that is the only option apart from the set $\mathcal{A}_0$, resulting in payoff for P arbitrarily close to $\underline{\pi}$
- If F^* puts all its weight on y's that maximize w, and $E_{F^*}[w](= \max w(y)) > V_A(w|\mathcal{A}_0)$, the same argument applies.

Proof of Lemma 2.2, contd.

- So suppose $E_{F^*}[w](= \max w(y)) = V_A(w|\mathcal{A}_0)$, implying existence of $(F, 0) \in \mathcal{A}_0$ which guarantees $\max w(y)$ to A
- Then under $\mathcal{A}_0$, A will select this technology; and will be willing to select it (and participate) if P deviates to a null contract, so $V_P(w) < V_P(0)$ and w could not be essential

Proof of Lemma 2.2, contd.

- So suppose $E_{F^*}[w](=\max w(y)) = V_A(w|\mathcal{A}_0)$, implying existence of $(F, 0) \in \mathcal{A}_0$ which guarantees $\max w(y)$ to A
- Then under $\mathcal{A}_0$, A will select this technology; and will be willing to select it (and participate) if P deviates to a null contract, so $V_P(w) < V_P(0)$ and w could not be essential
- For the last part, if it were not true we would have $E_{F^*}[w] > V_A(w|\mathcal{A}_0)$, we could find $\tilde{F}$ close to F^* (shift some weight to zero output) which would be chosen by A and result in worse profits for P (since $V_P(w) > 0$)

Separating Hyperplane argument

- Lemma 2.2 implies: *there does not exist* $F \in \Delta(Y)$ *satisfying*

$$E_F[w(y)] > V_A(w|\mathcal{A}_0) \quad \text{and} \quad E_F[y - w(y)] < V_P(w)$$

- i.e., S the convex hull of $(w(y), y - w(y))$, and of $T \equiv \{(u, v) \in \Re^2 | u > V_A(w|\mathcal{A}_0), v < V_P(w)\}$ do not have any interior points in common

Separating Hyperplane argument

- Lemma 2.2 implies: *there does not exist $F \in \Delta(Y)$ satisfying*

$$E_F[w(y)] > V_A(w|\mathcal{A}_0) \quad \text{and} \quad E_F[y - w(y)] < V_P(w)$$

- i.e., S the convex hull of $(w(y), y - w(y))$, and of $T \equiv \{(u, v) \in \Re^2 | u > V_A(w|\mathcal{A}_0), v < V_P(w)\}$ do not have any interior points in common

- So there exists k, λ, μ with $(\lambda, \mu) \neq (0, 0)$ such that

$$\begin{aligned} \mu[y - w(y)] &\geq k + \lambda w(y), \forall y \in Y \\ \mu V_P(w) &\geq k + \lambda V_A(w|\mathcal{A}_0) \end{aligned}$$

Separating Hyperplane argument, contd.

- Not difficult to check that $(\lambda, \mu) >> (0, 0)$, so we can normalize by setting $\mu = 1$
- $(V_P(w), V_A(w|\mathcal{A}_0))$ belongs to the boundary of both S and T, so $V_P(w) = k + \lambda V_A(w|\mathcal{A}_0)$
- So we get **Lemma 2.4:** for any non-zero eligible contract w:

$$\begin{aligned} y - w(y) &\geq k + \lambda w(y), \forall y \in Y \\ V_P(w) &= k + \lambda V_A(w|\mathcal{A}_0) \end{aligned}$$

Last Steps

- Consider the affine contract $w'(y) = \frac{y-k}{1+\lambda} \geq w(y), \forall y$ (using first $\geq$ in L2.4)
- Implies: $V_A(w'|\mathcal{A}_0) \geq V_A(w|\mathcal{A}_0)$
- P's worst case profit cannot go down, because if A chooses F' in response to w':

$$
\begin{aligned}
E_{F'}[y - w'(y)] &\geq k + \lambda E_{F'}[w'(y)] \\
&\geq k + \lambda V_A(w'|\mathcal{A}_0) \\
&\geq k + \lambda V_A(w|\mathcal{A}_0) = V_P(w)
\end{aligned}
$$

Last Steps

- Consider the affine contract $w'(y) = \frac{y-k}{1+\lambda} \geq w(y), \forall y$ (using first $\geq$ in L2.4)
- Implies: $V_A(w'|\mathcal{A}_0) \geq V_A(w|\mathcal{A}_0)$
- P's worst case profit cannot go down, because if A chooses F' in response to w':

$$\begin{aligned} E_{F'}[y - w'(y)] &\geq k + \lambda E_{F'}[w'(y)] \\ &\geq k + \lambda V_A(w'|\mathcal{A}_0) \\ &\geq k + \lambda V_A(w|\mathcal{A}_0) = V_P(w) \end{aligned}$$

- Finally, observe that any affine contract is dominated by corresponding linear contract (lump sum payment is non-negative, does not affect A's behavior, so can be eliminated)