ECON 2060 Contract Theory: Notes

Richard Holden
Harvard University
Littauer 225
Cambridge MA 02138
rholden@harvard.edu

September 6, 2016

Contents

1 Introduction

1.1 Situating Contract Theory

Think of (at least) three types of modeling environments

1. Competitive Markets: Large number of players → General Equilibrium Theory
2. Strategic Situations: Small number of players → Game Theory
3. Small numbers with design → Contract Theory & Mechanism Design

- Don't take the game as given
- Tools for understanding institutions

1.2 Types of Questions

1.2.1 Insurance

- 2 parties A & B
- A faces risk - say over income $Y_A = 0, 100, 200$ with probabilities $1/3, 1/3, 1/3$ and is risk-averse
- B is risk-neutral
- Gains from trade
- If A had all the bargaining power the risk-sharing contract is B pays A 100
- But we don't usually see full insurance in the real world
    1. Moral Hazard (A can influence the probabilities)
    2. Adverse Selection (There is a population of A's with different probabilities & only they know their type)

1.2.2 Borrowing & Lending

- 2 players
- A has a project, B has money
- Gains from trade
- Say return is $f(e, \theta)$ where e is effort and θ is the state of the world
- B only sees f not e or θ
- Residual claimancy doesn't work because of limited liability (could also have risk-aversion)
- No way to avoid the risk-return trade-off

1.2.3 Relationship Specific Investments

- A is an electricity generating plant (which is movable *pre hoc*)
- B is a coal mine (immovable)
- If A locates close to B (to save transportation costs) they make themselves vulnerable
- Say plant costs 100
- "Tomorrow" revenue is 180 if they get coal, 0 otherwise
- B's cost of supply is 20
- Zero interest rate
- NPV is 180-20-100=60
- Say the parties were naive and just went into period 2 cold
- Simple Nash Bargaining leads to a price of 100
- $\pi_A = (180 - 100) - 100 = -20$
- An illustration of the **Hold-Up Problem**
- Could write a long-term contract: bounded between 20 and 80 due to zero profit prices for A & B, maybe it would be 50
- But what is contract are incomplete – the optimal contract may be closer to no contract than a very fully specified one
- Maybe they should merge?

2 Mechanism Design

- Often, individual preferences need to be aggregated
- But if preferences are private information then individuals must be relied upon to reveal their preferences
- What constraints does this place on social decisions?
- Applications:
  - Voting procedures
  - Design of public institutions
  - Writing of contracts
  - Auctions
  - Market design
  - Matching

2.1 The Basic Problem

- Suppose there are I agents
- Agents make a collective decision x from a choice set X
- Each agent privately observes a preference parameter $\theta_i \in \Theta_i$
- Bernoulli utility function $u_i(x, \theta_i)$
- Ordinal preference relation over elements of $X \succcurlyeq_i (\theta_i)$
- Assume that agents have a common prior over the distribution of types
  - (i.e. the density $\phi(\cdot)$ of types on support $\Theta = \Theta_1 \times ... \times \Theta_I$ is common knowledge)

Remark 1. *The common prior assumption is sometimes referred to as the Harsanyi Doctrine. There is much debate about it, and it does rule out some interesting phenomena. However, it usefully rules out "betting pathologies" where participants can profitably bet against one another because of differences in beliefs.*

- Everything is common knowledge except each agent's own draw

Definition 1. *A Social Choice Function is a map $f : \Theta \to X$.*

Definition 2. *We say that f is Ex Post Efficient if there does not exist a profile $(\theta_1, ..., \theta_I)$ in which there exists any $x \in X$ such that $u_i(x, \theta_i) \geq u_i(f(\theta), \theta_i)$ for every i with at least one inequality strict.*

- ie. the SCF selects an alternative which is Pareto optimal given the utility functions of the agents
- There are multiple ways in which a social choice function ("SCF") might be implemented
  - Directly: ask each agent her type
  - Indirectly: agents could interaction through an institution or *mechanism* with particular rules attached
    - eg. an auction which allocates a single good to the person who announces the highest price and requires them to pay the price of the second-highest bidder (a second-price sealed bid auction).
- Need to consider both direct and indirect ways to implement SCFs

Definition 3. *A Mechanism $\Gamma = (S_1, ..., S_I, g(\cdot))$ is an $I + 1$ tuple consisting of a strategy set S_i for each player i and a function $g : S_1 \times ... \times S_I \to X$.*

- We'll sometimes refer to g as the "outcome function"
- A mechanism plus a type space $(\Theta_1, ..., \Theta_I)$ plus a prior distribution plus payoff functions $u_1, ..., u_I$ constitute a game of incomplete information. Call this game $\mathcal{G}$

Remark 2. *This is a normal form representation. At the end of the course we will consider using an extensive form when we study subgame perfect implementation.*

- In a first-price sealed-bid auction $S_i = \mathbb{R}_+$ and given bids $b_1, ..., b_I$ the outcome function $g(b_1, ..., b_I) = \left(\{y_i(b_1, ..., b_I)\}_{i=1}^I, \{t_i(b_1, ..., b_I)\}_{i=1}^I \right)$ such that $y_i(b_1, ..., b_I) = 1$ iff $i = \min\{j : b_j = \max\{b_1, ..., b_I\}\}$ and $t_i(b_1, ..., b_I) = -b_i y_i(b_1, ..., b_I)$

Definition 4. *A strategy for player i is a function $s_i : \Theta_i \to S_i$.*

Definition 5. *The mechanism Γ is said to Implement a SCF f if there exists equilibrium strategies $(s_1^*(\theta_1), ..., s_I^*(\theta_I))$ of the game $\mathcal{G}$ such that $g(s_1^*(\theta_1), ..., s_I^*(\theta_I)) = f(\theta_1, ..., \theta_I)$ for all $(\theta_1, ..., \theta_I) \in \Theta_1 \times ... \times \Theta_I$.*

- Loosely speaking: there's an equilibrium of $\mathcal{G}$ which yields the same outcomes as the SCF f *for all possible profiles of types.*
- We want it to be true no matter what the actual types (ie. draws) are

Remark 3. *We are requiring only* ***an*** *equilibrium, not a unique equilibrium.*

Remark 4. *We have not specified a solution concept for the game. The literature has focused on two solution concepts in particular: dominant strategy equilibrium and Bayes Nash equilibrium.*

- The set of all possible mechanisms is enormous!
- The *Revelation Principle* provides conditions under which there is no loss of generality in restricting attention to direct mechanisms in which agents truthfully reveal their types in equilibrium.

Definition 6. *A Direct Revelation Mechanism is a mechanism in which $S_i = \Theta_i$ for all i and $g(\theta) = f(\theta)$ for all $\theta \in (\Theta_1 \times ... \times \Theta_I)$.*

Definition 7. *The SCF f is Incentive Compatible if the direct revelation mechanism Γ has an equilibrium $(s_1^*(\theta_1), ..., s_I^*(\theta_I))$ in which $s_i^*(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and all i.*

2.2 Dominant Strategy Implementation

- A strategy for a player is weakly dominant if it gives her at least as high a payoff as any other strategy for all strategies of all opponents.

Definition 8. *A mechanism Γ Implements the SCF f in dominant strategies if there exists a dominant strategy equilibrium of Γ, $s^*(\cdot) = (s_1^*(\cdot), ..., s_I^*(\cdot))$ such that $g(s^*(\theta)) = f(\theta)$ for all $\theta \in \Theta$.*

- A strong notion, but a robust one
  - eg. don't need to worry about higher order beliefs
  - Doesn't matter if agents miscalculate the conditional distribution of types
  - Works for any prior distribution $\phi(\cdot)$ so the mechanism designer doesn't need to know this distribution

Definition 9. *The SCF f is Truthfully Implementable in Dominant Strategies if $s_i^*(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and $i = 1, ..., I$ is a dominant strategy equilibrium of the direct revelation mechanism $\Gamma = (\Theta_1, ..., \Theta_I, f(\cdot))$, ie*

$$u_i(f(\theta_i, \theta_{-i}), \theta_i) \geq u_i\left(f\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) \text{ for all } \hat{\theta}_i \in \Theta_i \text{ and } \theta_{-i} \in \Theta_{-i}. \quad (1)$$

Remark 5. *This is sometimes referred to as being "dominant strategy incentive compatible" or "strategy-proof".*

Remark 6. *The fact that we can restrict attention **without loss of generality** to whether $f(\cdot)$ in incentive compatible is known as the Revelation Principle (for dominant strategies).*

- This is very helpful because instead of searching over a very large space we only have to check each of the inequalities in (1).
  - Although we will see that this can be complicated (eg. when there are an uncountably infinite number of them).

Theorem 1. *(Revelation Principle for Dominant Strategies) Suppose there exists a mechanism Γ that implements the SCF f in dominant strategies. Then f is incentive compatible.*

Proof. The fact that Γ implements f in dominant strategies implies that there exists $s^*(\cdot) = (s_1^*(\cdot), ..., s_I^*(\cdot))$ such that $g(s^*(\theta)) = f(\theta)$ for all θ and that, for all i and $\theta_i \in \Theta_i$, we have

$$u_i(g(s_i^*(\theta_i), s_{-i}), \theta_i) \geq u_i(g(\hat{s}_i(\theta_i), s_{-i}), \theta_i) \text{ for all } \hat{s}_i \in S_i, s_{-i} \in S_{-i}.$$

In particular, this means that for all i and $\theta_i \in \Theta_i$

$$u_i\left(g\left(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})\right), \theta_i\right) \geq u_i\left(g\left(s_i^*\left(\hat{\theta}_i\right), s_{-i}^*(\theta_{-i})\right), \theta_i\right),$$

for all $\hat{\theta}_i \in \Theta_i, \theta_{-i} \in \Theta_{-i}$. Since $g(s^*(\theta)) = f(\theta)$ for all θ, the above inequality implies that for all i and $\theta_i \in \Theta_i$

$$u_i(f(\theta_i, \theta_{-i}), \theta_i) \geq u_i\left(f\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) \text{ for all } \hat{\theta}_i \in \Theta_i, \theta_{-i} \in \Theta_{-i},$$

which is precisely incentive compatibility. □

- Intuition: suppose there is an indirect mechanism which implements f in dominant strategies and where agent i plays strategy $s_i^*(\theta_i)$ when she is type θ_i. Now suppose we asked each agent her type and played $s_i^*(\theta_i)$ on her behalf. Since it was a dominant strategy it must be that she will truthfully announce her type.

2.2.1 The Gibbard-Satterthwaite Theorem

Notation 1. *Let $\mathcal{P}$ be the set of all rational preference relations $\succcurlyeq$ on X where there is no indifference*

Notation 2. *Agent i's set of possible ordinal preference relations on X are denoted $\mathcal{R}_i = \{\succcurlyeq_i : \succcurlyeq_i = \succcurlyeq_i(\theta_i) \text{ for some } \theta_i \in \Theta_i\}$*

Notation 3. *Let $f(\Theta) = (x \in X : f(\theta) = x \text{ for some } \theta \in \Theta)$ be the image of $f(\cdot)$.*

Definition 10. *The SCF f is Dictatorial if there exists an agent i such that for all $\theta \in \Theta$ we have:*

$$f(\theta) \in \{x \in X : u_i(x_i, \theta_i) \geq u_i(y, \theta_i), \forall y \in X\}.$$

- Loosely: there is some agent who always gets her most preferred alternative under f.

Theorem 2. *(Gibbard-Satterthwaite) Suppose: (i) X is finite and contains at least three elements, (ii) $\mathcal{R}_i = \mathcal{P}$ for all i, and (iii) $f(\Theta) = X$. Then the SCF f is dominant strategy implementable if and only if f is dictatorial.*

Remark 7. *Key assumptions are that individual preferences have unlimited domain and that the SCF takes all values in X.*

- The idea of a proof is the following: identify the pivotal voter and then show that she is a dictator
  - See Benoit (Econ Lett, 2000) proof
  - Very similar to Geanakoplos (Cowles, 1995) proof of Arrow's Impossibility Theorem
  - See Reny paper on the relationship
- This is a somewhat depressing conclusion: for a wide class of problems dominant strategy implementation is not possible unless the SCF is dictatorial
- It's a theorem, so there are only two things to do:
  - Weaken the notion of equilibrium (eg. focus on Bayes Nash equilibrium)
  - Consider more restricted environments
- We begin by focusing on the latter

2.2.2 Quasi-Linear Preferences

- An alternative from the social choice set is now a vector $x = (k, t_1, ..., t_I)$, where $k \in K$ (with K finite) is a choice of "project".
- $t_i \in \mathbb{R}$ is a monetary transfer to agent i
- Agent i's preferences are represented by the utility function

  $$u_i(x, \theta) = v_i(k, \theta_i) + (\bar{m}_i + t_i),$$

  where $\bar{m}_i$ is her endowment of money.
- Assume no outside parties
- Set of alternatives is:

  $$X = \left\{(k, t_1, ..., t_I) : k \in K, t_i \in \mathbb{R} \text{ for all } i \text{ and } \sum_i t_i \leq 0\right\}.$$

- Now consider the following mechanism: agent i receives a transfer which depends on how her announcement of type affects the other agent's payoffs through the choice of project. Specifically, agent i's transfer is exactly the externality that she imposes on the other agents.

- A SCF is ex post efficient in this environment if and only if:

$$\sum_{i=1}^{I} v_i\left(k\left(\theta\right), \theta_i\right) \geq \sum_{i=1}^{I} v_i\left(k, \theta_i\right) \text{ for all } k \in K, \theta \in \Theta, k\left(\theta\right).$$

Proposition 1. *Let $k^*(\cdot)$ be a function which is ex post efficient. The SCF $f = (k^*(\cdot), t_1, ..., t_I)$ is truthfully implementable in dominant strategies if, for all $i = 1, ..., I$*

$$t_i\left(\theta\right) = \left(\sum_{j \neq i} v_j\left(k^*\left(\theta\right), \theta_j\right)\right) + h_i\left(\theta_{-i}\right), \tag{2}$$

where h_i is an arbitrary function.

- This is known as a Groves-Clarke mechanism

Remark 8. *Technically this is actually a Groves mechanism after Groves (1973). Clarke (1971) discovered a special case of it where the transfer made by an agent is equal to the externality imposed on other agent's if she is pivotal, and zero otherwise.*

- Groves-Clarke type mechanisms are implementable in a quasi-linear environment

- Are these the only such mechanisms which are?

- Green and Laffont (1979) provide conditions under which this question is answered in the affirmative

- Let $\mathcal{V}$ be the set of all functions $v : K \to \mathbb{R}$

Theorem 3. *(Green and Laffont, 1979) Suppose that for each agent $i = 1, ..., I$ we have $\{v_i(\cdot, \theta_i) : \theta_i \in \Theta_i\} = \mathcal{V}$. Then a SCF $f = (k^*(\cdot), t_1(\cdot), ..., t_I(\cdot))$ in which $k^*(\cdot)$ satisfies*

$$\sum_{i=1}^{I} v_i\left(k\left(\theta\right), \theta_i\right) \geq \sum_{i=1}^{I} v_i\left(k, \theta_i\right),$$

for all $k \in K$ (efficient project choice) is truthfully implementable in dominant strategies only if $t_i(\cdot)$ satisfies (2) for all $i = 1, ..., I$.

- ie. if every possible valuation function from K to $\mathbb{R}$ arises for some type then a SCF which is truthfully implementable must be done so through a mechanism in the Groves class

- So far we have focused on only one aspect of ex post efficient efficiency–that the efficient project be chosen

- Another requirement is that none of the numeraire be wasted

- The condition is sometimes referred to as "budget balance" and requires

$$\sum_i t_i(\theta) = 0 \text{ for all } \theta \in \Theta.$$

- Can we satisfy both requirements?
- Green and Laffont (1979) provide conditions under which this question is answered in the negative

Theorem 4. *(Green and Laffont, 1979) Suppose that for each agent $i = 1, ..., I$ we have $\{v_i(\cdot, \theta_i) : \theta_i \in \Theta_i\} = \mathcal{V}$. Then there does not exists a SCF $f = (k^*(\cdot), t_1(\cdot), ..., t_I(\cdot))$ in which $k^*(\cdot)$ satisfies*

$$\sum_{i=1}^{I} v_i(k(\theta), \theta_i) \geq \sum_{i=1}^{I} v_i(k, \theta_i),$$

for all $k \in K$ (efficient project choice) and

$$\sum_i t_i(\theta) = 0 \textit{ for all } \theta \in \Theta,$$

(budget balance).

- Either have to waste some of the numeraire or give up on efficient project choice
- Can get around this if there is one agent whose preferences are known
  - Maybe one agent doesn't care about project choice
  - eg. the seller in an auction
  - Maybe the project only affects a subset of the population...
- Need to set the transfer for the "no private information" type to $t_{BB}(\theta) = -\sum_{i \neq 0} t_i(\theta)$ for all θ.
- This agent is sometime referred to as the "budget breaker"
- We will return to this theme later in the course (stay tuned for Legros-Matthews)

2.3 Bayesian Implementation

- Now move from dominant strategy equilibrium as the solution concept to Bayes-Nash equilibrium
- A strategy profile implements an SCF f in Bayes-Nash equilibrium if for all i and all $\theta_i \in \Theta_i$ we have

$$E_{\theta_{-i}}\left[u_i\left(g\left(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})\right), \theta_i\right)|\theta_i\right] \geq$$
$$E_{\theta_{-i}}\left[u_i\left(g\left(\hat{s}_i(\theta_i), s_{-i}^*(\theta_{-i})\right), \theta_i\right)|\theta_i\right],$$

  for all $\hat{s}_i \in S_i$.

- Again, we are able to make use of the revelation principle
- Same logic as in dominant strategy case
  - If an agent is optimizing by choosing $s_i^*(\theta_i)$ in some mechanism Γ then if we introduce an intermediary who will play that strategy for her then telling the truth is optimal conditional on other agents doing so. So truth telling is a (Bayes-Nash) equilibrium of the direct revelation game (ie. the one with the intermediary).

Remark 9. *Bayesian implementation is a weaker notion than dominant strategy implementation. Every dominant strategy equilibrium is a Bayes-Nash equilibrium but the converse is false. So any SCF which is implementable in dominant strategies can be implemented in Bayes-Nash equilibrium, but not the converse.*

Remark 10. *Bayesian implementation requires that truth telling give the highest payoff* ***averaging*** *over all possible types of other agents. Dominant strategy implementation requires that truth telling be best for* ***every*** *possible type of other agent.*

- Can this relaxation help us overcome the negative results of dominant strategy implementation
- Again consider a quasi-linear environment
- Under the conditions of Green-Laffont we couldn't implement a SCF truthfully and have efficient project choice and budget balance
- Can we do better in Bayes-Nash?
- A direct revelation mechanism known as the "expected externality mechanism" due to d'Aspremont and Gérard-Varet (1979) and Arrow (1979) answers this in the affirmative
- Under this mechanism the transfers are given by:

$$t_i(\theta) = E_{\tilde{\theta}_{-i}}\left[\sum_{j\neq i} v_j\left(k^*\left(\theta_i, \tilde{\theta}_{-i}\right), \tilde{\theta}_j\right)\right] + h_i(\theta_{-i}).$$

- The first term is the expected benefit of other agents when agent i announces her type to be θ_i and the other agents are telling the truth

2.4 Participation Constraints

- So far we have worried a lot about incentive compatibility
- But we have been assuming that agents have to participate in the mechanism
- What happens if participation is voluntary?

2.4.1 Public Project Example

- Decision to do a project or not $K = \{0, 1\}$
- Two agents with $\Theta_i = \{L, H\}$ being the (real-valued) valuations of the project
- Assume that $H > 2L > 0$
- Cost of project is $c \in (2L, H)$
- An ex post efficient SCF has $k^* (\theta_1, \theta_2) = 1$ if either $\theta_1 = H$ or $\theta_2 = H$ and $k^* (\theta_1, \theta_2) = 0$ if (and only if) $\theta_1 = \theta_2 = L$
- With no participation constraint we can implement this SCF in dominant strategies using a Groves scheme
- By voluntary participation we mean that an agent can withdraw at any time (and if so, does not get any of the benefits of the project)
- With voluntary participation agent 1 must have $t_1 (L, H) \geq -L$
  - Can't have to pay more than L when she values the project at L because won't participate voluntarily
- Suppose both agents announce H. For truth telling to be a dominant strategy we need:

$$\begin{aligned} Hk^* (H, H) + t_1(H, H) &\geq Hk^* (L, H) + t_1(L, H) \\ H + t_1(H, H) &\geq H + t_1(L, H) \\ t_1(H, H) &\geq t_1(L, H) \end{aligned}$$

- But we know that $t_1 (L, H) \geq -L$, so $t_1(H, H) \geq -L$
- Symmetrically, $t_2(H, H) \geq -L$
- So $t_1 (L, H) + t_2(H, H) \geq -2L$
- But since $c > 2L$ we can't satisfy $t_1 (L, H) + t_2(H, H) \geq -c$
- Budget breaker doesn't help either, because $t_{BB} (\theta_1, \theta_2) \geq 0$ for all (θ_1, θ_2) and hence $t_0 (H, H) \geq 0$ and we can't satisfy

$$t_0(H, H) + t_1(H, H) + t_2(H, H) \leq -c.$$

2.4.2 Types of Participation Constraints

- Distinguish between three different types of participation constraint depending on timing (of when agents can opt out of the mechanism)
- Ex ante: before the agents learn their types, ie:

$$U_i(f) \geq E_{\theta_i} [\bar{u}_i (\theta_i)] . \tag{3}$$

- Interim: after agents know their own types but before the take actions (under the mechanism), ie:

$$U_i(\theta|f) = E_{\theta_{-i}}\left[u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right)|\theta_i\right] \geq \bar{u}_i\left(\theta_i\right) \quad \text{for all } \theta_i. \tag{4}$$

- Ex post: after types have been announced and an outcome has been chosen (it's a direct revelation mechanism)

$$u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \geq \bar{u}_i\left(\theta_i\right) \quad \text{for all } \left(\theta_i, \theta_{-i}\right) \tag{5}$$

- A question of when agents can agree to be bound by the mechanism
- Constraints are most severe when agents can withdraw ex post and least severe when they can withdraw ex ante. This can be seen from the fact that (5)⇒(4)⇒(3) but the converse doesn't hold

Theorem 5. *(Myerson-Satterthwaite) Suppose there is a risk-neutral seller and risk-neutral buyer of an indivisible good and suppose their respective valuations are drawn from* $[\underline{\theta}_1, \bar{\theta}_1] \in \mathbb{R}$ *and* $[\underline{\theta}_2, \bar{\theta}_2] \in \mathbb{R}$ *according to strictly positive densities with* $(\underline{\theta}_1, \bar{\theta}_1) \cap (\underline{\theta}_2, \bar{\theta}_2) \neq \varnothing$. *Then there does not exist a Bayesian incentive compatible SCF which is ex post efficient and gives every type non-negative expected gains from participation.*

- Whenever gains from trade are possible but not certain there is no ex post efficient SCF which is incentive compatible and satisfies interim participation constraints

Remark 11. *This applies to all voluntary trading institutions, including all bargaining processes.*

3 Adverse Selection (Hidden Information)

3.1 Static Screening

3.1.1 Introduction

- A good reference for further reading is Fudenberg & Tirole chapter 7
- Different to "normal" Adverse Selection because 1 on 1, not a market setting
- 2 players: Principal and the Agent
- Payoff: Agent $G\left(u\left(q, \theta\right) - T\right)$, Principal $H\left(v\left(q, \theta\right) + T\right)$ where $G\left(\cdot\right), H\left(\cdot\right)$ are concave functions and q is some verifiable outcome (eg. output), T is a transfer, θ is the Agent's private information
- Don't use the concave transforms for now
- Say Principal is a monopolistic seller and the Agent is a consumer
- Let $v(q, \theta) = -cq$

- Principal's payoff is $T - cq$ where T is total payment (pq)
- $u(q, \theta) = \theta V(q)$
- Agent's payoff is $\theta V(q) - T$ where $V(\cdot)$ is strictly concave
- θ is type (higher $\theta \rightarrow$ more benefit from consumption)
- $\theta = \theta_1, ..., \theta_n$ with probabilities $p_1, ..., p_n$
- Principal only knows the distribution of types
- Note: relationship to non-linear pricing literature
- Assume that the Principal has all the bargaining power
- Start by looking at the first-best outcome (ie. under symmetric information)

First Best Case I: Ex ante no-one knows θ, ex post θ is verifiable

- Principal solves

$$\max_{(q_i,T_i)} {}^{n}_{i=1} p_i(T_i - cq_i)$$

$$s.t. {}^{n}_{i=1} p_i(\theta_i V(q_i) - T_i) \geq \overline{U} \qquad \text{(PC)}$$

First Best Case II: Ex ante both know θ

- Normalize $\overline{U}$ to 0
- Principal solves

$$\max_{(q_i,T_i)} \{T_i - cq_i\}$$

$$s.t. \theta_i V(q_i) - T \geq 0 \qquad \text{(PC)}$$

- The PC will bind, so $T_i = \theta_i V(q_i)$
- So they just solve $\max_{q_i} \{\theta_i V(q_i) - cq_i\}$
- FOC $\theta_i V'(q_i) = c$
- This is just perfect price discrimination – efficient but the consumer does badly
- Case I folds into II by offering a contingent contract

Second-Best

- Agent knows θ_i but the Principal doesn't
- First ask if we can achieve/sustain the first best outcome
- ie. will they naturally reveal their type
- say the type is θ_2
- if they reveal themselves their payoff is $\theta_2 V(q_2^*) - T_2^* = 0$
- if they pretend to be θ_1 their payoff is $\theta_2 V(q_2^*) - T_1^* = \theta_2 V(q_1^*) - \theta_1 V(q_1^*) = (\theta_2 - \theta_1) V(q_1^*) > 0$ since $\theta_2 > \theta_1$
- can't get the first-best

Second-best with n types

- First to really look at this was Mirrlees in his 1971 optimal income tax paper – normative
- Positive work by Akerlof, Spence, Stiglitz
- Revelation Principle very useful: can look at / restrict attention to contracts where people reveal their true type *in equilibrium*
- Without the revelation principle we would have the following problem for the principal

$$\max_{T(q)} \{ {}_{i=1}^{n} p_i (T(q_i) - cq_i) \}$$
$$\text{subject to}$$
$$\theta_i V(q_i) - T(q_i)) \geq 0, \forall i \qquad \text{(PC)}$$
$$q_i = \arg\max_q \{ \theta_i V(q) - T(q)) \}, \forall i \qquad \text{(IC)}$$

- But the revelation principle means that there is no loss of generality in restricting attention to optimal equilibrium choices by the buyers
- We can thus write the Principal's Problem as

$$\max_{(q_i, T_i)} \{ {}_{i=1}^{n} p_i (T_i - cq_i) \}$$
$$\text{subject to}$$
$$\theta_i V(q_i) - T_i) \geq 0, \forall i \qquad \text{(PC)}$$
$$\theta_i V(q_i) - T_i \geq \theta_i V(q_j) - T_j, \forall i, j \qquad \text{(IC)}$$

- Incentive compatibility means the Agent truthfully reveals herself
- This helps a lot because searching over a schedule $T(q)$ is hard
- Before proceeding with the n types case return to a two type situation

Second-best with* 2 *types

- Too many constraints to be tractable (there are $n(n-1)$ constraints of who could pretend to be whom)
- 2 types with $\theta_H > \theta_L$
- Problem is the following:

$$\max\left\{p_H(T_H - cq_H) + p_L(T_L - cq_L)\right\}$$
$$s.t.(i)\ \theta_H V(q_H) - T_H \geq \theta_H V(q_L) - T_L \quad \text{(IC)}$$
$$(ii)\ \theta_L V(q_L) - T_L \geq 0 \quad \text{(PC)}$$

- We have eliminated two constraints: the IC constraint for the low type and the PC constraint for the high type
- Why was this ok?
- The low type constraint must be the only binding PC (high types can "hide behind" low types)
- And the low type won't pretend to be the high type
- PC must bind otherwise we could raise T_L and the Principal will always be happy to do that
- IC must always bind otherwise the Principal could raise T_H (without equality the high type's PC would not bind) – also good for the Principal
- So $\theta_H V(q_H) - T_H = \theta_H V(q_L) - T_L$ and $\theta_L V(q_L) - T_L = 0$
- Now substitute to get an unconstrained problem:

$$\max_{q_L, q_H}\left\{p_H\left(\theta_H V(q_H) - \theta_H V(q_L) + \theta_L V(q_L) - cq_H\right) + p_L\left(\theta_L V(q_L) - cq_L\right)\right\}$$

- The FOCs are
$$p_H\theta_H V'(q_H) - p_H c = 0$$
and
$$p_L\theta_L V'(q_L) - p_L c + p_H\theta_L V'(q_L) - p_H\theta_H V'(q_L) = 0$$
- The first of these simplifies to $\theta_H V'(q_H) = c$ (so the high type chooses the socially efficient amount)
- The second of these simplifies to the following:

$$\theta_L V'(q_L) = \frac{c}{1 - \frac{1-p_L}{p_L}\frac{\theta_H - \theta_L}{\theta_L}} > c$$

(so the low type chooses too little)

- $q_H = q_H^*$ and $q_L < q_L^*$

- No incentive reason for distorting q_H because the low type isn't pretending to be the high type
- But you do want to discourage the high type from pretending to be the low type – and hence you distort q_L
- We can check the IC constraint is satisfied for the low type

$$\theta_H V(q_H) - T_H = \theta_H V(q_L) - T_L \text{ (high type's IC is binding)}$$

now recall that (recalling that $\theta_H > \theta_L, q_H > q_L$), so we have

$$\theta_L V(q_L) - T_L \geq \theta_L V(q_H) - T_H$$

- So the low type's IC is satisfied
- High type earns rents – PC does not bind
- Lots of applications: optimal taxation, banking, credit rationing, implicit labor contracts, insurance, regulation (see Bolton-Dewatripont for exposition)

3.1.2 Optimal Income Tax

- Mirrlees (Restud, 1971)
- Production function $q = \mu e$ (for each individual), where q is output, μ is ability and e is effort
- Individual knows μ and e but society does not
- Distribution of μs in the population, μ_L and μ_H in proportions π and $1-\pi$ respectively
- Utility function $U(q - T - \psi(e))$ where T is tax (subsidy if negative) and $\psi(e)$ is cost of effort (presumably increasing and convex)
- The government's budget constraint is $\pi T_L + (1-\pi)T_H \geq 0$
- Veil of Ignorance – rules are set up before the individuals know their type
- So the first-best problem is:

$$\max_{e_L, e_H, T_L, T_H} \{\pi U(\mu_L e_L - T_L - \psi(e_L)) + (1-\pi) U(\mu_H e_H - T_H - \psi(e_H))\}$$
$$\text{subject to}$$
$$\pi T_L + (1-\pi) T_H \geq 0$$

- But the budget constraint obviously binds and hence $\pi T_L + (1-\pi)T_H = 0$
- Then we have $T_H = -\pi T_L / (1-\pi)$

- The maximization problem can be rewritten as

$$\max_{e_L,e_H,T_L} \{\pi U\left(\mu_L e_L - T_L - \psi(e_L)\right) + (1-\pi)U\left(\mu_H e_H + (\pi T_L/1-\pi) - \psi(e_H)\right)\}$$

- The FOCs are

(i) $-U'(\mu_L e_L - T_L - \psi(e_L)) = U'\left(\mu_H e_H + (\pi T_L/1-\pi) - \psi(e_H)\right)$

(ii) $\mu_L = \psi'(e_L)$

(iii) $\mu_H = \psi'(e_H)$

- Choose e_L, e_H efficiently in the first-best
- Everyone has same marginal cost of effort so the higher marginal product types work harder
- (i) just says the marginal utilities are equated
- Hence $\mu_L e_L - T_L - \psi(e_L) = \mu_H e_H + T_H - \psi(e_H)$
- The net payoffs are identical so you are indifferent between which type you are
- Consistent with Veil of Ignorance setup
- There is no DWL because of the lump sum aspect of the transfer

Second-Best

- Could we sustain the first-best?
- No because the high type will pretend to be the low type, $\mu_H e = q_L$ so $q_L - T_L - \psi\left(q_L/\mu_H\right) > q_L - T_L - \psi\left(e_L\right)$ since $q_L/\mu_H < e_L$
- Basically the high type can afford to slack because they are more productive - hence no self sustaining first-best
- The Second-Best problem is

$$\max_{e_L,e_H,T_L,T_H} \{\pi U\left(\mu_L e_L - T_L - \psi(e_L)\right) + (1-\pi)U\left(\mu_H e_H - T_H - \psi(e_H)\right)\}$$
$$s.t.(i)\mu_H e_H - T_H - \psi(e_H) \geq \mu_L e_L - T_L - \psi(\mu_L e_L/\mu_H)$$
$$(ii)\pi T_L + (1-\pi)T_H \geq 0$$

- Solving yields $e_H = e_H^*$
- and $\mu_L = \psi'(e_L) + \beta(1-\pi)\left(\mu_L - \mu_L/\mu_H \psi'(\mu_L e_L/\mu_H)\right)$
- where $\beta = \frac{U_L' - U_H'}{U_L'}$ (marginal utilities evaluated at their consumptions levels)
- but $U_L < U_H$ so $U_L' > U_H'$ (by concavity) and hence $0 < \beta < 1$

- Since $\psi(\cdot)$ is convex we have $\psi'\left(\frac{\mu_L e_L}{\mu_H}\right) < \psi'(e_L)$
- $\mu_L > \psi'(e_L) + \beta(1-\pi)(\mu_L - \mu_L/\mu_H \psi'(e_L))$
- and hence:
$$\psi'(e_L) < \frac{\mu_L - \beta(1-\pi)\mu_L}{1-\beta(1-\pi)\mu_L/\mu_H} < \mu_L$$
- (the low type works too little)
- To stop the high type from misrepresenting themselves we have to lower the low type's required effort and therefore subsidy
- High type is better off $\rightarrow$ lose the egalitarianism we had before for incentive reasons
- Can offer a menu (q_L, T_L), (q_H, T_H) and people self select
- If you have a continuum of types there would be a tax schedule $T(q)$
- Marginal tax rate of the high type is zero (because they work efficiently) so $T'(q) = 0$ at the very top and $T'(q) > 0$ elsewhere with a continuum of types

3.1.3 Regulation

- Baron & Myerson (Ecta, 1982)
- The regulator/government is ignorant but the firm knows its type
- Firm's characteristic is $\beta \in \{\underline{\beta}, \overline{\beta}\}$ with probabilities ν_1 and $1-\nu_1$
- Cost is $c = \beta - e$
- Cost is verifiable
- Cost of effort is $\psi(e) = e^2/2$
- Let $\Delta\beta = \overline{\beta} - \underline{\beta}$ and assume $\Delta\beta < 1$
- Government wants a good produced with the lowest possible subsidy - wants to minimize expected payments to the firm
- The First-Best is simply
$$\min_e \{\beta - e + e^2/2\}$$
- The FOC is $e^* = 1$ and the firm gets paid $\beta - 1/2$
- Can we sustain the FB?
- No because $p_L = \beta_L - 1/2$ and $p_H = \beta_H - 1/2$

Second-Best

- Two cost levels $\underline{c}$ and $\overline{c}$
- Two price levels $\underline{p}$ and $\overline{p}$ (payments)
- Government solves

$$\min \left\{ \nu_1 \underline{p} + (1-\nu_1)\overline{p} \right\}$$
$$s.t.(i)\ \underline{p} - \underline{c} - e^2/2 \geq \overline{p} - \overline{c} - (\overline{e} - \Delta\beta)^2/2$$
$$(ii)\ \overline{p} - \overline{c} - \overline{e}^2/2 \geq 0$$

- noting that $\underline{e} = \overline{e} - \Delta\beta$ (from cost equation and low type pretending to be high type)
- Define $\underline{s} = \underline{p} - \underline{c} = \underline{p} - \underline{\beta} + \underline{e}$ and $\overline{s} = \overline{p} - \overline{c} = \overline{p} - \overline{\beta} + \overline{e}$ (these are the "subsidies")
- The government's problem is now

$$\min \left\{ \nu_1 \left(\underline{s} + \underline{\beta} - \underline{e} \right) + (1-\nu_1)\overline{s} + \overline{\beta} - \overline{e} \right\}$$
$$s.t.(i)\ \underline{s} - e^2/2 \geq \overline{s} - (\overline{e} - \Delta\beta)^2/2$$
$$(ii)\ \overline{s} - \overline{e}^2/2 \geq 0$$

- Since the constraints must hold with equality we can substitute and write this as an unconstrained problem

$$\min_{\underline{e},\overline{e}} \left\{ \nu_1 \left(\frac{\overline{e}^2}{2} + \underline{e}^2/2 - \frac{(\overline{e} - \Delta\beta)^2}{2} \right) + (1-\nu_1) \left(\frac{\overline{e}^2}{2} - \overline{e} \right) \right\}$$

- The FOCs are

  (1) $\underline{e} = 1$

  (2) $\nu_1 \overline{e} - \nu_1 (\overline{e} - \Delta\beta) + (1-\nu_1)\,\overline{e} - (1-\nu_1) = 0$
- (2) implies that:
$$\overline{e} = \frac{1 - \nu_1 - \nu_1\Delta\beta}{1-\nu_1} = 1 - \frac{\nu_1\Delta\beta}{1-\nu_1}$$
- The low cost ("efficient") type chooses $\underline{e} = 1$
- The high cost ("bad") types chooses $\overline{e} = 1 - \frac{\nu_1\Delta\beta}{1-\nu_1}$
- Offer a menu of contracts: fixed price or a cost-sharing arrangement
- The low cost firm takes the fixed price contract, becomes the residual claimant and then chooses the efficient amount of effort
- See also Laffont & Tirole (JPE, 1986) – costs observable

3.1.4 The General Case – n types and a continnum of types

- Problem of all the incentive compatibility constraints
- It turns out that we can replace the IC constraints with downward adjacent types
- The constraints are then just:

  (i) $\theta_i V(q_i) - T_i \geq \theta_i V(q_{i-1}) - T_{i-1} \quad \forall i = 2, ..., n$

  (ii) $q_i \geq q_{i-1} \quad \forall i = 2, ..., n$

  (iii) $\theta V(q_1) - T_1 \geq 0$

- (ii) is a monotonicity condition
- It is mathematically convenient to work with a continuum of types – and we will
- Let $F(\theta)$ be a cdf and $f(\theta)$ the associated density function on the support $[\underline{\theta}, \overline{\theta}]$
- The menu being offered is $T(\theta), q(\theta)$
- The problem is

$$\max_{T(\cdot),q(\cdot)} \left\{ \int_{\underline{\theta}}^{\overline{\theta}} [T(\theta) - cq(\theta)] f(\theta) d\theta \right\}$$

$$s.t.(i)\ \theta V(q(\theta) - T(\theta) \geq \theta V\left(q\left(\widehat{\theta}\right)\right) - T\left(\widehat{\theta}\right) \forall \theta, \widehat{\theta} \quad \text{(IC)}$$

$$(ii)\ \theta V(q(\theta) - T(\theta) \geq 0, \forall \theta \quad \text{(PC)}$$

- We will be able to replace all the IC constraints with a Local Adjacency condition and a Monotonoicity condition

Definition 11. *An allocation $T(\theta), q(\theta)$ is implementable if and only if it satisfies IC $\forall \theta, \widehat{\theta}$*

Proposition 2. *An allocation $T(\theta), q(\theta)$ is implementable if and only if $\theta V'(q(\theta)) \frac{dq(\theta)}{d\theta} - T'(\theta) = 0$ (the local adjacency condition) and $\frac{dq(\theta)}{d\theta} \geq 0$ (the monotonicity condition).*

Proof. $\Rightarrow$ direction:
Let $\widehat{\theta} = \arg\max_{\theta} \left\{ \theta V\left(q\left(\widehat{\theta}\right)\right) - T\left(\widehat{\theta}\right) \right\}$. Now $\frac{d}{d\widehat{\theta}} = \theta V'\left(q\left(\widehat{\theta}\right)\right) - \frac{dq(\widehat{\theta})}{d\widehat{\theta}} - T'\left(\widehat{\theta}\right)$
so $\theta V'(q(\theta)) - \frac{dq(\theta)}{d\theta} - T'(\theta) = 0, \forall \theta$
Now, by revealed preference:

$$\theta V(q(\theta)) - T(\theta) \geq \theta V(q(\theta')) - T(\theta')$$

and

$$\theta' V(q(\theta')) - T(\theta') \geq \theta' V(q(\theta)) - T(\theta)$$

combining these yields:

$$\theta\left[V\left(q\left(\theta\right)\right)-V\left(q\left(\theta'\right)\right)\right]\geq T\left(\theta\right)-T'\left(\theta\right)\geq\theta'\left[V\left(q\left(\theta\right)\right)-V\left(q\left(\theta'\right)\right)\right]$$

the far RHS can be expressed as $\left(\theta-\theta'\right)\left(V\left(q\left(\theta\right)\right)-V\left(q\left(\theta'\right)\right)\right)\geq 0$
hence if $\theta>\theta'$ then $q\left(\theta\right)\geq q\left(\theta'\right)$ □

- This really just stems from the **Single-Crossing Property** (or **Spence-Mirrlees Condition**), namely $\frac{\partial U}{\partial q}$ is increasing in θ
- Note that this is satisfied with the separable functional form we have been using–but need not be satisfied in general
- Higher types are "even more prepared" to buy some increment than a lower type

Proof. $\Leftarrow$ direction

Let $W\left(\theta,\widehat{\theta}\right)=\theta V\left(q\left(\widehat{\theta}\right)\right)-T\left(\widehat{\theta}\right)$. Fix θ and suppose the contrary. This implies that $\exists\widehat{\theta}$ such that $W\left(\theta,\widehat{\theta}\right)>W\left(\theta,\theta\right)$.

Case 1: $\widehat{\theta}>\theta$

$$W\left(\theta,\widehat{\theta}\right)-W\left(\theta,\theta\right)=\int_{\theta}^{\widehat{\theta}}\frac{\partial W}{\partial\tau}\left(\theta,\tau\right)d\tau=\int_{\theta}^{\widehat{\theta}}\theta V'\left(q\left(\tau\right)\right)\frac{dq}{d\tau}-T'\left(\tau\right)d\tau$$

But $\tau>\theta$ implies that:

$$\begin{aligned}&\int_{\theta}^{\widehat{\theta}}\theta V'\left(q\left(\tau\right)\right)\frac{dq}{d\tau}-T'\left(\tau\right)d\tau\\ \leq\;&\int_{\theta}^{\widehat{\theta}}\left(\tau V'\left(q\left(\tau\right)\right)\frac{dq}{d\tau}-T'\left(\tau\right)\right)d\tau=0\end{aligned}$$

because the integrand is zero. Contradiction. Case 2 is analogous. □

- This proves that the IC constraints are satisfied globally, not just the SOCs (the common error)
- Now we write the problem as:

$$\max_{T(\cdot),q(\cdot)}\left\{\int_{\underline{\theta}}^{\overline{\theta}}\left[T\left(\theta\right)-cq\left(\theta\right)\right]f(\theta)d\theta\right\}$$

$$s.t.(i)\ \theta V'(q\left(\theta\right))\frac{dq\left(\theta\right)}{d\theta}-T'\left(\theta\right)\geq 0\quad\forall\theta \qquad \text{(Local Adjacency)}$$

$$(ii)\ \frac{dq\left(\theta\right)}{d\theta}\geq 0\quad\forall\theta \qquad \text{(Monotonicity)}$$

$$(iii)\underline{\theta}V\left(q\left(\underline{\theta}\right)\right)-T\left(\underline{\theta}\right)=0 \qquad \text{(PC-L)}$$

- Let $W\left(\theta\right)\equiv W\left(\theta,\theta\right)=\theta V\left(q\left(\theta\right)\right)-T\left(\theta\right)=\max_{\hat{\theta}}\left\{\theta V(q(\hat{\theta}))-T\left(\hat{\theta}\right)\right\}$

- Recall that in the 2 type case we used the PC for the lowest type and the IC for the other type
- We could have kept on going for higher and higher types
- Now, from the FOCs:

$$\frac{dW(\theta)}{d\theta} = \theta V'(q(\theta))\frac{dq}{d\theta} - \frac{dT}{d\theta} + V(q(\theta)) = V(q(\theta))$$

  (by adding $V(q(\theta))$ to both sides)

$$W(\theta) - W(\underline{\theta}) = \int_{\underline{\theta}}^{\theta} \frac{dW(\tau)}{d\tau} d\tau = \int_{\underline{\theta}}^{\theta} V(q(\tau))\, d\tau$$

(change of measure trick)

- But $W(\underline{\theta}) = 0$ (PC of low type binding at the optimum)
- Now $T(\theta) = -\int_{\underline{\theta}}^{\theta} V(q(\tau))\, d\tau + \theta V(q(\theta))$ (by substitution)
- So the problem is now just

$$\max_{q(\cdot)} \left\{ \int_{\underline{\theta}}^{\overline{\theta}} \left[\theta V(q(\theta)) - \int_{\underline{\theta}}^{\theta} V(q(\tau))\, d\tau - cq(\theta) \right] f(\theta)\, d\theta \right\}$$

$$s.t. \frac{dq(\theta)}{d\theta} \geq 0 \quad \forall \theta$$

- Proceed by ignoring the constraint for the moment and tackle the double integral using integration by parts
- Recall that

$$\int_{\underline{\theta}}^{\overline{\theta}} uv' = uv\Big|_{\underline{\theta}}^{\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} u'v$$

- So let $v' = f(\theta)$ and $u = \int V(q(\tau))\, d\tau$, and we then have

$$\begin{aligned}
\int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\underline{\theta}}^{\theta} V(q(\tau))\, d\tau \right] f(\theta)\, d\theta &= \int_{\underline{\theta}}^{\theta} V(q(\tau))\, d\tau F(\theta) \Big|_{\underline{\theta}}^{\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} V(q(\theta)) F(\theta)\, d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}} V(q(\tau))\, d\tau - \int_{\underline{\theta}}^{\overline{\theta}} V(q(\theta)) F(\theta)\, d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}} V(q(\theta)) [1 - F(\theta)]\, d\theta
\end{aligned}$$

- So we can write the problem as:

$$\max_{q(\cdot)} \left\{ \int_{\underline{\theta}}^{\overline{\theta}} ((\theta V(q(\theta) - cq(\theta)) f(\theta) - V(q(\theta)[1 - F(\theta)])\, d\theta \right\}$$

- Now we can just do pointwise maximization (maximize under the integral for all values of θ)

$$\theta V'(q(\theta)) = V'(q(\theta))\left(\frac{1-F(\theta)}{f(\theta)}\right) + c, \quad \forall\theta \tag{6}$$

- From 6 we can say the following:
  (1)
  $$\theta = \overline{\theta} \to \overline{\theta} V\left(q\left(\overline{\theta}\right)\right) = c$$

  (2)
  $$\theta < \overline{\theta} \to \overline{\theta} V\left(q\left(\overline{\theta}\right)\right) > c$$

  ($q(\theta)$ is too low)

- Since efficiency requires $\theta V'(q(\theta)) = c$
- Now differentiate (6) and solve for $\frac{dq}{d\theta} \geq 0$
- This implies that $\frac{f(\theta)}{1-F(\theta)}$ is increasing in θ (*this is a sufficient condition in general, but is a necessary and sufficient condition in this buyer-seller problem*)
- This property is known as the **Monotone Hazard Rate Property**
- It is satisfied for all log-concave distributions
- We've been considering the circumstance where θ announces their type, θ^a and gets a quantity $q(\theta^a)$ and pays a tariff of $T(\theta^a)$
- This can be reinterpreted as: given $\widehat{T}(q)$, pick q
- For each q there can only be one $T(q)$ by incentive compatibility
- $\widehat{T}(q) = T(\theta^{-1}(q))$
- The optimization problem becomes

$$\max_q \left\{\theta V(q) - \widehat{T}(q)\right\}$$

- The FOC is $\theta V'(q) = \widehat{T}'(q) \equiv p(q)$

$$p(q) = \frac{p(q(\theta))}{\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right) + c$$

$$\frac{p-c}{p} = \frac{1-F(\theta)}{\theta f(\theta)}$$

- Recall that we ignored the constraint $\frac{dq}{d\theta} \geq 0$

- The FOC implies
$$\left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) V'\left(q\left(\theta\right)\right) = c$$
- Differentiating this wrt θ yields
$$\frac{dq}{d\theta} = -\frac{g'\left(\theta\right) v'\left(q\left(\theta\right)\right)}{v''\left(q\left(\theta\right)\right) g\left(\theta\right)},$$
where $g\left(\theta\right) = \theta - \frac{1-F(\theta)}{f(\theta)}$
- Since the following holds
$$\theta V'(q(\theta)) = V'(q(\theta)) \left(\frac{1-F(\theta)}{f(\theta)}\right) + c$$
we have
$$V'(q(\theta)) = \frac{c}{\theta - \left[\left(1-F(\theta)\right)/f(\theta)\right]}$$
- We require that $V'(q(\theta))$ be falling in θ and hence require that $\theta - \frac{1-F(\theta)}{f(\theta)}$ be increasing in θ
- That is, that the hazard rate be increasing
- Now turn attention to $T(q)$
- $\widehat{T'}(q) > c$ except for at the very top where $\widehat{T'} = c$
- Therefore it can't be convex
- Note that
$$1 - \frac{c}{p} = \frac{1-F\left(\theta\right)}{\theta f\left(\theta\right)}$$
$$\frac{\theta f(\theta)}{1-F(\theta)} \uparrow \theta \Leftrightarrow \frac{dp}{dq} < 0$$
- And note that $\frac{dp}{dq} = \widehat{T''}(q)$
- So the IHRC $\Rightarrow \frac{dp}{dq} < 0$
- If the IHRC does not hold the Monotonicity Constraint binds and we need to applying "Ironing" (See Bolton & Dewatripont)
- Use Pontryagin's Principle to find the optimal cutoff points
- Require $\lambda(\theta_1) = \lambda(\theta_2) = 0$, where λ is the Lagrange multiplier
- Still get optimality and the top and sub-optimality elsewhere

3.1.5 Random Schemes

- Key paper is Maskin & Riley (RAND, 1984)
- A deterministic scheme is always optimal if the seller's program is convex
- But if the ICs are such that the set of incentive feasible allocations is not convex then random schemes may be superior

dtbpF3.3529in2.0678in0ptFigure

- Both types are risk-averse
- So S loses money on the low type, but may be able to charge enough more to the high type to avoid the randomness if the high type is more risk-averse
- If they are sufficiently more risk-averse (ie. the types are far enough apart), then the random scheme dominates
- Say: announce $\theta = \theta^a$ and get a draw from a distribution, so get $(\widetilde{q}, \widetilde{T})$
- If the high type is less risk-averse than the low type then the deterministic contract dominates
  - The only incentive constraints that matter are the downward ones
  - So if the high type is less risk-averse then S loses money on that type from introducing randomness
  - And doesn't gain anything on the low type, because her IR constraint is already binding and so can't extract more rents from her

3.1.6 Extensions and Applications

- Jullien (2000) and Rochet & Stole (2002) consider more general PCs (egs. type dependent or random)
- Classic credit rationing application: Stiglitz & Weiss (1981)

Multi-Dimensional Types

- So far we have assumed that a single parameter θ captures all relevant information
- Laffont-Maskin-Rochet (1987) were the first to look at this
- They show that "bunching" is more likely to occur in a two-type case than a one-type case (ie. Monotone Hazard Rate condition violated)
- Armstrong (Ecta, 1996) provides a complete characterization
  - Shows that some agents are always excluded from the market at the optimum (unlike the one-dimensional case where there is no exclusion)
  - In one dimension if the seller increases the tariff uniformly by ε then profits go up by ε on all types whose IR was slack enough (so that they still participate), but lose on all the others

  - With multi-dimensional types the probability that an agent had a surplus lower than ε is a higher order term in ε – so the loss is lower from the increase even if there is exclusion

- Rochet-Chone (1997) shows that
  - Upward incentive constraints can be binding at the optimum
  - Stochastic constracts can be optimal
  - There is no generalization of the MHRC which can rule out bunching

- Armstrong (1997) shows that with a large number of independently valued dimensions the the optimal contract can be approximated by a two-part tariff

Aside: Multi-Dimensional Optimal Income Taxation

- Mirrlees (JPubE, 1976) considered the problem of multi-dimensional optimal income taxation
- Strictly harder than the above problems because he doesn't assume quasi-linear utility functions only
- He shows how, when $m < n$ (i.e. the number of characteristics is smaller than the number of commodities), the problem can be reduced to a single eliptic equation which can be solved by well-known method
- When $m \geq n$ (i.e. the number of characteristics is at least at large than the number of commodities) the above approach does not lead to a single second-order partial differential equation, but a system of m second-order partial differential equations for the m functions M_j
- Numerical evidence has shown recently that a lot of the conclusions from the one-dimensional case go away in multiple dimensions (eg. the no distortion at the top result)
- But the system of second-order PDEs seem very hard to solve

3.2 Dynamic Screening

- Going to focus on the situation where there are repeated interactions between an informed and uninformed party
- We will assume that the informed party's type is fixed / doesn't change over time
  - There is a class of models where the agent gets a new draw from the distribution of types each period (see BD §9.2 for details)
- The main new issue which arises is that there is (gradual) elimination of the information asymmetry over time
- Renegotiation a major theme
  - Parties may be subject to a contract, but can't prevent Pareto improving (and therefore mutual) changes to the contract

3.2.1 Durable good monopoly

- There is a risk-neutral seller ("S") and a risk-neutral buyer ("B")
- Normalize S's cost to zero
- B's valuation is $\overline{b}$ or $\underline{b}$ with probabilities $\mu, 1-\mu$ and assume that $\overline{b} > \underline{b} > 0$
  - This is common knowledge
- B knows their valuation, S does not
- Trade-off is $\underline{b}$ vs. $\mu_1 \overline{b}$ and assume that $\mu_1 \overline{b} > \underline{b}$ (i.e. low types are not very likely)
- 2 periods
- Assume that the good is a durable good and there is a discount factor of δ which is common to B and S

Commitment

- Assume that S can commit not to make any further offers
- Under this assumption it can be shown that the Revelation Principle applies
- Contract: if B announces $\overline{b}$ then with probability $\overline{x_1}$ B gets the good today and with probability $\overline{x_2}$ they get the good tomorrow. B pays $\overline{p}$ for this
- Similarly for $\underline{b} \to \underline{x_1}, \underline{x_2}, \underline{p}$
- S solves:

$$\max_{\overline{x_1}, \overline{x_2}, \underline{x_1}, \underline{x_2}} \left\{ \mu_1 \overline{p} + (1-\mu_1)\underline{p} \right\}$$

$$s.t.(i) \quad \overline{b}\left(\overline{x_1}(1+\delta) + (1-\overline{x_1})\overline{x_2}\delta\right) - \overline{p} \geq \overline{b}\left(\underline{x_1}(1+\delta) + (1-\underline{x_1})\underline{x_2}\delta\right) - \underline{p}$$

$$(ii) \quad \underline{b}\left[\underline{x_1}(1+\delta) + (1-\underline{x_1})\underline{x_2}\delta\right] - \underline{p} \geq 0$$

- In fact, both constraints will hold with equality
- Let

$$\overline{X_1} = \overline{x_1}(1+\delta) + (1-\overline{x_1})\overline{x_2}\delta$$

$$\underline{X_1} = \underline{x_1}(1+\delta) + (1-\underline{x_1})\underline{x_2}\delta$$

- $\underline{p} = \underline{b}\underline{X_1}$ and $\overline{p} = \overline{bX_1} - \overline{b}\underline{X_1} + \underline{b}\underline{X_1}$ since the high type's IC constraint and low type's IR constraint both bind at the optimum
- So:

$$\max \left\{ \mu_1 \left[\overline{bX_1} - \overline{b}\underline{X_1} + \underline{b}\underline{X_1}\right] + (1-\mu_1)\underline{b}\underline{X_1} \right\}$$

$$s.t.(i) \quad 0 \leq \overline{X_1} \leq 1+\delta$$

$$(ii) \quad 0 \leq \underline{X_1} \leq 1+\delta$$

- The constraints are just physical constraints
- Notice that the coefficient on $\underline{X_1}$ is $\underline{b} - \mu_1 \overline{b} < 0$
- Similarly for $\overline{X_1} : \mu_1 \overline{b} > 0$
- Conclusion: since $\mu_1 \overline{b} > \underline{b}$ by assumption it is optimal to set $\overline{X_1} = 1 + \delta, \underline{X_1} = 0$ and $\underline{p} = 0$, $\overline{p} = \overline{b} + \delta \overline{b}$ (ie. what it's worth to the high type)
- Just a repetition of the one period model (S faces a stationary problem because of commitment)

No Commitment

- Now consider the case where S cannot commit
- Suppose S can't commit and date 1 not to make further offers in period 2
- Study the Perfect Bayesian Equilibria ("PBE") of the game
- Basically, S has the following choices
  - (1) Sell to both types at date 1
  - (2) Sell to both types at date 2
  - (3) Never sell to the low type
- Under (1) $\underline{p} = \underline{b} + \delta \underline{b}$, $\Pi_1 = \underline{b} + \delta \underline{b}$
- Under (2) $p_2 = \underline{b}$, $p_1 = \overline{b} + \delta \underline{b}$ since $\overline{b} + \delta \overline{b} - p_1 = \delta(\overline{b} - \underline{b})$, by incentive compatibility
- Notice that under (2) $\Pi_2 = \mu_1 \left(\overline{b} + \delta \underline{b} \right) + (1 - \mu_1) \delta \underline{b} = \mu_1 \overline{b} + \delta \underline{b}$
- Hence $\Pi_2 > \Pi_1$ since $\mu_1 \overline{b} > \underline{b}$
- Now consider strategy (3) – only sell to the high type in both periods
- Under this strategy $p_1 = \overline{b} + \delta \overline{b}$, $p_2 = \overline{b}$
- Need to credibly commit to keep the price high in period 2
- The high type buys with probability ρ_1 in period 1 and $1 - \rho_1$ in period 2
  - No pure strategy equilibrium because if $p_2 = \underline{b}$ then the high type doesn't want to buy in period 1 and if $p_2 = \overline{b}$ then high type wants to buy in period 1 so that can't be a continuation equilibrium
- Use Bayes' Rule to obtain:

$$\begin{aligned} pr[\overline{b} \quad | \quad \text{declined first offer}] &= \frac{\mu_1(1-\rho_1)}{\mu_1(1-\rho_1) + (1-\mu_1)} \\ &= \frac{\mu_1(1-\rho_1)}{1-\mu_1\rho_1} = \sigma \end{aligned}$$

- Condition for the Seller to keep price high is:

$$\sigma \geq \underline{b}/\overline{b}$$

- Note: this is the Pareto efficient PBE
- If fact it will hold with equality (ρ_1 as high as possible), and can be written as:

$$\frac{\mu_1(1-\rho_1)}{1-\mu_1\rho_1} = \underline{b}/\overline{b}$$

- Early buyers are good, but can't have too many (in order to maintain credibility)
- Solving yields:

$$\rho_1^* = \frac{\mu_1\overline{b} - \underline{b}}{\mu_1\left(\overline{b} - \underline{b}\right)}$$

- Therefore the Seller's expected profit from strategy (3) is:

$$\begin{aligned} & \mu_1\rho_1\left(\overline{b} + \delta\overline{b}\right) + \mu_1(1-\rho_1)\delta\overline{b} \\ = & \mu_1\rho_1\overline{b} + \mu_1\delta\overline{b} \\ = & \mu_1\overline{b}\left[\frac{\mu_1\overline{b} - \underline{b}}{\mu_1\left(\overline{b} - \underline{b}\right)}\right] + \mu_1\delta\overline{b} \end{aligned}$$

- Expected profit from strategy (2) was $\mu_1\overline{b} + \delta\underline{b}$
- Strategy (3) is preferred to strategy (2) iff:

$$\begin{aligned} \mu_1 & > \frac{\overline{b}\underline{b}\,(1+\delta) - \delta\underline{b}}{\delta\overline{b}^2 - \delta\overline{b}\underline{b} + \overline{b}\underline{b}} \\ & \equiv \overline{\mu_2} \end{aligned}$$

- Check that $\overline{\mu_2} > \overline{\mu_1} = \underline{b}/\overline{b}$ (and it is)
- Now consider a T period model (Hart & Tirole, 1988)

$$\exists 0 \leq \overline{\mu_1} \leq \overline{\mu_2} < ... < \overline{\mu_T} < 1 \text{ such that}$$

- $\mu_1 < \overline{\mu_1} \Rightarrow$ sell to low types at date 1
- $\overline{\mu_2} > \mu_1 > \overline{\mu_1} \Rightarrow$ sell to low types at date 2
- $\overline{\mu_3} > \mu_1 > \overline{\mu_2} \Rightarrow$ sell to low types at date 3
- $\overline{\mu_T} > \mu_1 > \overline{\mu_{T-1}} \Rightarrow$ sell to low types at date T
- $\mu_1 > \overline{\mu_1} \Rightarrow$ never sell to low types
- In addition it can be shown that $\overline{\mu_i}$ is independent of T

- Also: $\overline{\mu_i}$ is weakly decreasing in δ $\forall i$ - if people are more patient the seller will do more screening
- Also: $\overline{\mu_i}$ has a well defined limit as $\delta \to 1$
- $\overline{\mu_i} \to 1$ as $T \to \infty$
- COASE CONJECTURE (Coase 1972): When periods become very short it's like $\delta \to 1$
- *As period length goes to zero bargaining is over (essentially) immediately – so the price is the value that the low type puts on it* $\Rightarrow$ the seller looses all their monopoly power

3.2.2 Non-Durable Goods

- Every period S can sell 1 or 0 units of a good to B
- Can think of this as renting the good
- B ends up revealing her type in a separating equilibrium
- Commitment solution is essentially the same as the Durable Good case
- Non-Commitment solution is very different
- S offers
$$r_1; r_2(Y), r_2(N); r_3(YY), r_3(YN), r_3(NY), r_3(NN); ...$$
- Consider Perfect Bayesian Equilibria ("PBE")
- Here the problem is that S can't commit not to be tough in future periods (people effectively reveal their type) – a *Ratcheting Problem*
- 2 period model: is ratcheting a problem?
- Say they try to implement the durable good solution:
$$\begin{aligned} S1 &: \underline{b} + \delta\underline{b} \\ S3 &: \overline{b} + \delta\overline{b}, \underline{b} \\ S2 &: p_1 = \overline{b} + \delta\underline{b}, p_2 = \underline{b} \end{aligned}$$
- in the service model $\widehat{p}_2(N) = \underline{b}, \widehat{p}_2(Y) = \overline{b} \Rightarrow \widehat{p}_1 = \overline{b}(1-\delta) + \delta\underline{b}$ since $\underline{b} - \widehat{p}_1 + \delta(\overline{b} - \overline{b}) = \delta(\overline{b} - \underline{b})$
- So ratcheting isn't a problem with 2 periods
- But this breaks down with many periods
- *Screening fails because the price you have to charge in period 1 to induce the high types to buy is below the price at which the low type is prepared to buy*
- Take T large and suppose that $\mu_{i-1} < \mu_1 < \overline{\mu_i}$

- Consider date $i-1$:

$$\bar{b} - r_{i-1} \geq (\bar{b} - \underline{b}) \left(\delta + \delta^2 + ...\right) \simeq (\bar{b} - \underline{b}) \frac{\delta}{1-\delta}$$

- if T is large, and this $\geq \bar{b} - \underline{b}$ if $\delta > \frac{1}{2}$
- $\Rightarrow r_{i-1} < \underline{b}$
- Now the low type will buy at $i-1$
- *Screening blows-up*

Proposition 3. *Assume $\delta > \frac{1}{2}$. Then for any prior beliefs μ_1 $\exists k$ such that $\forall T$ and $t < T - l, r_t = \underline{b}$*

- Non Coasian dynamics: pools for a while and then separates
- In the Durable Goods case: Coasian dynamics - separates for a while and then pools
- Can get around it with a contract (long-term contract)
- Consider a service model but allow S to offer long-term contracts
- But don't prevent them from lowering the price (to avoid this just becoming a commitment technology)
- Can offer "better" contracts
- This returns us to the Durable Goods case - the ratcheting disappears (see Hart and Tirole)
- *A long-term contract is just like a durable good*
- *As soon as you go away from commitment in dynamic models the Revelation Principle fails – the information seeps out slowly here*

3.2.3 Soft Budget Constraint

- Kornai (re Socialist Economies)
- Dewatripont & Maskin
- Government faces a population of firms each needing one unit of capital
- Two types of firms: α good, quick types - project gets completed and yields $Rg > 1$ (money) and Eg (private benefit to firm / manager). There are also $1-\alpha$ bad, slow types - no financial return, zero or negative private benefit, but can be refinanced at further cost $1 \rightarrow \Pi_b^*$ financial benefit and a private benefit of Eb $(1 < \Pi_b^* < 2)$
- Can the Government commit not to refinance?
- If yes then only the good types apply – and this is first-best

- If no then bad types also apply – and bad types are negative NPV so the outcome is sub-optimal
- Decentralization may aid commitment (free riding actually helps!)
- We will return to this idea when we study financial contracting
  - Dispersed creditors can act as a commitment no to renegotiate
- Transition in Eastern Europe (Poland and banking reform v. mass privatization)

3.2.4 Non Committment

- Go back to a regulation setting
- Firm has cost type $\theta \in \{\beta_L, \beta_H\}$ and let $\Delta\beta = \beta_H - \beta_L$
- τ indexes time periods, $\tau = 1, 2$
- Cost of production is $C_\tau = \beta - e_\tau$
- Cost of effort is $\psi(e_\tau)$
- Regulator pays subsidy t_τ
- Welfare is $W_\tau = S - (1+\lambda)(C_\tau + t_\tau) + U_\tau$, where $U_\tau = t_\tau - \psi(e_\tau)$
- Firm gets $t - \psi(e)$
- FB: $\psi'(e^*) = 1$
- Suppose government's prior is $\Pr(\beta = \beta_L) = \nu_1$
- Let (t_L, C_L) be the contracts chosen by type β_L (so that effort is $e_L = \beta_L - C_L$) and similarly for type β_H
- In a one period problem the solution is just like before

$$\begin{aligned}\psi'(e_L^*) &= 1 \\ \psi'(e_H^*) &< 1.\end{aligned}$$

- Two periods without committment–can't commit to second period contracts
- Move away from the Revelation Principle
- First consider a continuum of types
- $\beta \in [\underline{\beta}, \bar{\beta}]$ with prior CDF $F_1(\cdot)$
- Posterior formed according to Bayes Rule and denoted $F_2(\cdot)$
- Going to consider the equilibria of the game between the regulator and the firm
- Solution concept: PBE

- The regulator's strategy is an incentive scheme in each period $C_1 \to t_1(C_1)$ and $C_2 \to t_2(C_2; t_1(\cdot), C_1)$
- The firm may quit the relationship and earn zero payoff
- Firm's strategy is a participation choice and an effort level
- An equilibrium will be said to be fully revealing if $\beta \to C_1(\beta) = \beta - e_1(\beta)$ is one to one

Theorem 6 ((Laffont-Tirole (1988))). *For any first period incentive scheme $t_1(\cdot)$ there does not exist a fully separating continuation equilibrium*

- If a firm reveals itself then it earns zero rent in the second period because the FB obtains
- Consider a firm of type β and think about a deviation from the revealing equilm strategy of producing at a cost as if it was type β and instead produces as if its cost was $\beta + d\beta$
- By the envelope theorem it makes a second-order loss of profit from this deviation
- But it gets a first-order rent in period 2 because in this equilm the regulator believes it to be type $\beta + d\beta$
- Striking result–highlights the importance of the committment assumption
- Moreover, there is no non-degenerate sub-interval of $[\underline{\beta}, \bar{\beta}]$ over which separation occurs
- If there is "small uncertainty" (i.e. $\left|\underline{\beta} - \bar{\beta}\right|$ is small) then the equilibrium must involve "much pooling" in the sense that one can find two types which are arbitrarily far apart who pool
- A natural kind of equilibrium to look for is a *partition equilibrium* (like in Crawford-Sobel cheap talk)
- In such an equilm $[\underline{\beta}, \bar{\beta}]$ is partitioned into a countable number of ordered intervals such that in period one all types in that interval choose the same cost level
- Full pooling is a degenerate partition equilm where there is only on such sub-interval
- When does such an equilm exist?

Definition 12. *We say that ψ is α-convex if $\psi'' \geq \alpha$ everywhere*

- Necessary condition: suppose that ψ is α-convex and that the equilm is a partition equilm. If C^k and C^l are two equilm first-period cost levels then $\left|C^k - C^l\right| \geq \delta/\alpha$.
- i.e. the minimum distance between two equilm costs is equal to the discount factor divided by the curvature of the disutility of effort
- e.g. in the quadratic case $\psi(e) = \alpha e^2/2$

- Sufficient condition: suppose that ψ is α-convex and that the distribution is log-concave. Then if the regulator offers a finite set of cost transfer pairs $\left\{t^k, C^k\right\}$ such that $\left|C^k - C^l\right| \geq \delta/\alpha$ for all k, l then there exists a partition equilibrium of the continuation game
- Much more stringent assumptions than when sending a message is costless (as in Crawford-Sobel)
- Further results are available in the two-type case
- An issue of how many types of contracts are allowed–unlike the static or committment case there can be a benefit from offering more contracts than types
- Three possibilities in the two-contract case: (i) Efficient type's IC constraint only is binding, (ii) Inefficient type's IC constraint only is binding, (iii) Both IC constraints are binding
- Turns out that (iii) is a real possibility
- Ask the following: is the optimal static mechanism incentive compatible in period 1 in the non-committment case?
- In that allocation the efficient type's IC constraint binds and the inefficient type's IR constraint binds
- Not IC: the efficient type earns zero rent in period 2 by revealing herself, but a positive rent by not doing so
- To make it IC the regulator has to offer her a larger period 1 transfer, to compensate for the loss of period 2 surplus
- But if the discount factor is high enough then the inefficient type will choose this contract in period 1
- They are then believed to be the efficient type and have to produce a lot in period 2
- If we could force them to participate then this would be a deterrent
- But they can opt out and get zero in period 2
- This is called the "take the money and run" strategy
- To avoid this pooling can be better

Proposition 4. *There exists a $\bar{\delta} > 0$ such that for all $\delta \leq \bar{\delta}$ the separating equilibrium is preferred to the pooling equilibrium and for all $\delta > \bar{\delta}$ the pooling equilibrium is better than the separating equilibrium*

- A large discount factor is plausible–especially since there are just two periods here!
- By offering a single contract in period 1 the regulator is essentially committing not to update her beliefs and ensures the optimal static mechanism in period 2
- When period 2 matters a lot relative to period 1 then this is optimal

4 Persuasion and Information Transmission

4.1 Cheap Talk

- Crawford-Sobel (Ecta, 1982)
- Main question: how much information can be transmitted when communication is costless, but interests are not necessarily aligned?
- cf. signaling models: the key ingredient there is that communication is *costly*
  - eg. Spence job market signaling: to get separation need education to more costly for certain types than for others
- Two parties: a decision maker who is uninformed, and an informed expert
- DM is to make a decision $d \in [0, 1]$
- State of nature is $\theta \in [0, 1]$
- DM has a uniform prior about θ
- DM's payoff is $U(d, \theta) = -(d - \theta)^2$
- The E(xpert) knows the value of θ and her payoff is $V(b, \theta, d) = -(d - (\theta + b))^2$
- $b \geq 0$ is a measure of the bias of the expert
- E may send a message $m \in [0, 1]$
- Timing:

  1. 1. E observes θ
     2. E sends m to DM
     3. DM chooses d

- Solution concept: PBE

Proposition 5. *For all b there exists a "babbling equilbrium" in which E sends a random message ("babbles") and hence no information is conveyed.*

- Intuition: in a babbling equilibrium DM believes there is no information content in the message. E then has no incentive to send an informative message, so is happy to babble
- Bigger question: are there informative equilibria?
- Preliminary question: are there equilibria in which information is truthfully conveyed?

Proposition 6. *There exists an equilibrium in which information is fully revealed if and only if $b = 0$.*

- Proof sketch: suppose $b > 0$ and E truthfully revealed θ. In this equilibrium she is believed, but her payoff could be increased in some states by deviating to a message $\theta + b$–a contradiction.

- Now we constuct an equilibrium in which *some* information is conveyed
- Let DM's posterior distribution about the value of θ given m be $G(\theta|m)$
- Given quadratic preferences

$$\begin{aligned} d^*(m) &\equiv \max_{d \in [0,1]} \left\{ \int U(d,\theta)\, G(\theta|m)\, d\theta \right\} \\ &= E[\theta|m]. \end{aligned}$$

- E knows this, of course, and could be faced with the following problem
- Suppose message m leads to action d and message m' leads to action $d' > d$
- Also, suppose that in state $\theta'' > \theta'$ E prefers d' to d but in state θ' prefers d to d'
- Noting that V satisfies single crossing, $d^2V/d\theta dd > 0$ and hence E prefers d' to d for all $\theta > \theta''$
- Therefore, by the Intermediate Value Theorem, there exists a state $\hat{\theta}$ such that $\theta' < \hat{\theta} < \theta''$ in which E is indifferent between d and d'
- This is the same as saying that the distance between E's bliss point and d in state $\hat{\theta}$ is the same as the distance between the bliss point and d'
- ie. $\hat{\theta} + b - d = d' - \left(\hat{\theta} + b\right)$
- So E sends message m for all $\theta < \hat{\theta}$ and message m' for all $\theta > \hat{\theta}$
- For this to be an equilbrium we need to find d, d' and $\hat{\theta}$ such that

$$\begin{aligned} \hat{\theta} + b - d &= d' - \left(\hat{\theta} + b\right), \text{ and} \\ d(m) &= E[\theta|m]. \end{aligned}$$

- Solving we have

$$\begin{aligned} d &= \frac{\hat{\theta}}{2}, \\ d' &= \frac{1+\hat{\theta}}{2}. \end{aligned}$$

- Substituting into $\hat{\theta} + b - d = d' - \left(\hat{\theta} + b\right)$ we have

$$\hat{\theta} = \frac{1}{2} - 2b.$$

- Clearly such an equilibrium exists
- Moreover the cutoff $\hat{\theta}$ is uniquely determined by b

- If bias is too big then non-existence (ie. $b > 1/4$)
- This is a particular equilm with just two partitions
- But when bias is "small" there exist equilibria with more than two partitions

Theorem 7 (Crawford-Sobel)**.** *There exists a partition equilibrium of every size (ie. number of partitions) from 1 (completely uninformative) to* $N(b)$ *(the most informative).*

- Many equilibria!
- One thing to focus on is the impossibility of perfectly informative communication
- Another is the following quite general message: when preferences are somewhat aligned cheap talk "can" improve both party's payoff
- Cheap talk is just that: not an announcement in a mechanism, not a costly signal, just an unverifiable utterance
- One might think it could never help much (eg. Yogi Berra: "a verbal contract isn't worth the paper it's written on"), but the CS theorem shows that it *could*
- A large literature explores concrete settings in which it *does*
- Basic idea: cheap talk (by construction) does not directly affect payoffs, but it can affect them indirectly

4.2 Improved Communication

4.2.1 Conversation

- Krishna-Morgan (JET, 2004)
- Mulitiple messages from E can be subsumed as one message
- What about two-sided communication
- DM knows "nothing" and hence has no new information to reveal, but can act as a randomization device
- Illustration: suppose $b = 1/12$
- If only E gets to talk then the most informative equilm reveals whether θ is above or below 1/3
- Timing:

1. 1. E observes θ
   2. E and DM meet face to face
   3. E delivers a "report"
   4. DM chooses d

- Consider the following equim:

- In the meeting E reveals whether $\theta > 1/6$ or not and sends some other message (this determines whether the meeting is a "success" or not
- If E sends the message that $\theta \leq 1/6$ then the meeting is deemed a failure and DM chooses $d = 1/12$
- If E sends $\theta > 1/6$ then the report is conditional on the success or failure of the meeting
- If the meeting was a failure then $d = 7/12$ (the optimal action conditional on $\theta > 1/6$
- But if the meeting was a "success" then the report further partitions the interval $[1/6, 1]$ into $[1/6, 5/12]$ and $[5/12, 1]$
- In the first subinterval $d^* = 7/24$ and in the second $d^* = 17/24$
- If $\theta = 1/6$ then E prefers $d = 1/12$ to $d = 7/12$
- So we need "uncertainty" about the outcome of the meeting–otherwise E would not be willing to reveal whether the state was above or below 1/6
- If $\theta < 1/6$ then E would say $\theta \in [1/6, 5/12]$ and induce $d = 7/24$ and if $\theta > 1/6$ then E would announce $\theta < 1/6$ and induce $d = 1/12$ rather than $d = 7/12$
- It turns out that when $\theta = 1/6$ then with probability $p = 16/21$ E is indifferent between $d = 1/12$ and the lottery where she gets $d = 7/12$ with probability p and $d = 7/24$ with probability $1 - p$ gets $d = 7/12$
- When $\theta < 1/6$ E prefers $d = 1/12$ to the lottery and when $\theta > 1/6$ E prefers the lottery
- So can we get the meeting to be successful with probability $p = 16/21$?
- KM show that we can, as follows
- Suppose E sends message (low, A_i) or $(high, A_i)$ and DM sends a message A_j with $i, j \in \{1, ..., 21\}$
- Low means $\theta \leq 1/6$ and high means $\theta > 1/6$
- The A_i and A_j parts of the message serve as a coordination device about the success of the meeting
- E chooses A_i randomly (i.e. from a uniform distribution)
- DM does similarly for A_j
- The meeting is deemed a success if $0 \leq i - j < 16$ or of $j - 1 > 5$ and a failure otherwise
- With this structure the probability that the meeting is a success is exactly $p = 16/21$
- So more information is conveyed than in any CS equilm
- Striking thing: having the DM participate in the conversation helps even though she is completely uninformed
- Aumann and Hart (2003) show that even with unlimited communication full revelation is impossible (cf. Geanakoplos-Polemarcharkis)

4.2.2 Delegation

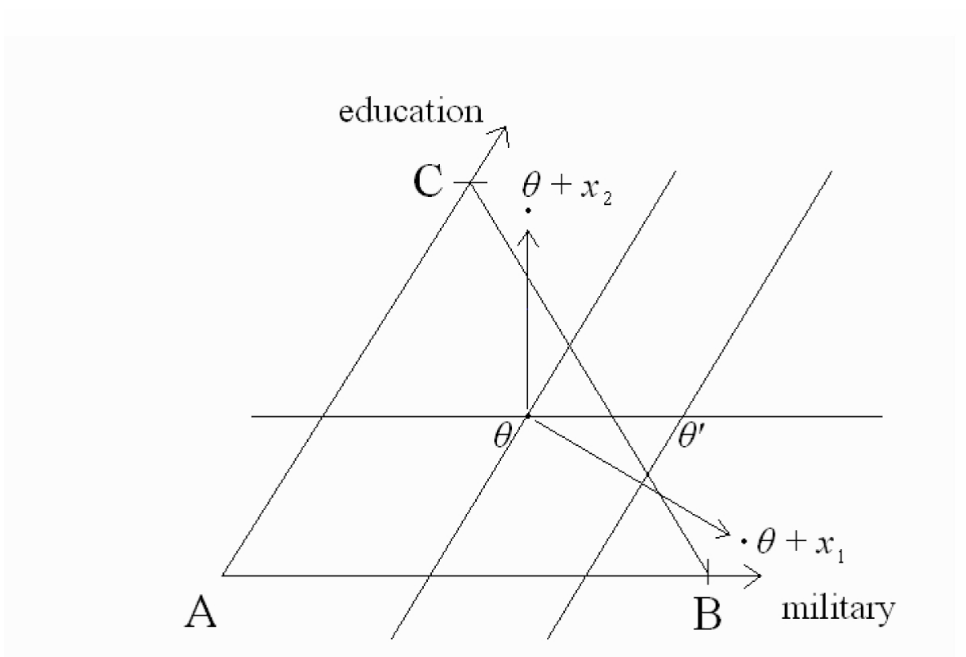
- Can we do better by delegating to E?
- Tradeoff: E has her own preferences and is thus biased, but she is also informed
- Suppose $b = 1/12$ then direct computation yields a payoff of $-1/36$ in the most informative partition equilm of the CS model, but under delegation the action which is chosen is $d = \theta + b$, by construction, and the payoff is $-b^2 = -1/144$, so delegation is optimal
- This conclusion is more general than this example (see Dessein, 2002)
- DM can do even better by cambining the amount of delegation/discretion
- Here the optimal thing to do is limit E's discretion to $d \in [0, 1-b]$

4.2.3 Compensation

- An obvious ommision in what we did is to preclude the possibility of compensating E for her advice
- Can we do better with an optimal contract?
- Now add a transfer such that the payoffs are
  - DM's payoff is $U(d, \theta) = -(d-\theta)^2 - t$
  - E's payoff is $V(b, \theta, d) = -(d-(\theta+b))^2 + t$
- Again use mechanism design to find the optimal contract
- Can apply the revelation principle here and restrict attention to mechanisms/contracts whereby E announces d and θ truthfully in equilm
- Aside: this isn't cheap talk any more–talk affects payoffs directly here
- Suppose $t\left(\hat{\theta}\right) = 2b\left(1-\hat{\theta}\right)$ and $d\left(\hat{\theta}\right) = \hat{\theta}$, then the FB decision is acheived and there is full revelation
- But this is costly for DM
- eg. when $b = 1/12$ her payoff is $-1/12$, whereas it is $-1/36$ in the best CS equilm
- General result: Krishna-Morgan (2004): Full revelation is in general feasible, but never optimal

4.2.4 Multiple Senders and Multiple Dimensions

- Battaglini (2002): two sender cheap talk with a one dimensional state space
  - Also showed that with a multi-dimensional Euclidean state space a perfectly revealing PBE can be constructed
  - Moreover, there are no out of equilm messages and so these equilbria survive any refinements which place restrictions on out of equilm beliefs
  - Construction: each sender conveys information only along directions in which her interests coincide with DM (ie. directions which are orthogonal to the bias of E)
  - Since these generically span the whole state space DM can extract all the information and perfectly identify the true state
- Ambrus-Takahashi (2007) consider restricted state spaces
  - eg. some policies may not be feasible
  - or some may never be chosen by DM (and so they are not rationalizable)
- AT provide the following example



- Suppose DM needs to allocate a fixed budget to "education," "military spending," and "healthcare"
- Suppose there are two perfectly informed experts, a left-wing E and a right-wing E
- Left-wing E has a bias towards spending more on education, while right-wing E has a bias towards spending more on the military, but both of them are unbiased with respect to healthcare

- The state space in this example is represented by triangle ABC
- At B it is optimal for DM to spend the whole budget on the military
- At C it is optimal to spend all money on education
- At A it is optimal to spend no money on either education or military
- Left-wing E's bias is orthogonal to AB in the direction of C and right-wing E's bias is orthogonal to AC in the direction of B
- Battaglini's solution would have left-wing E report along a line parallel to AC (like asking how much money to spend on the military), right-wing E to report along a line parallel to AB
- But here it is not true that any pair of such reports identifies a point in the state space!
- Look at state θ
- If left-wing E sends a truthful report, then the right-wing analyst can send reports that put you outside the state space
- ie. they say that expenditure should be larger than the budget
- Doesn't happen in equilm, but have to specify out of equilm beliefs and this can cause problems for the construction
- Key points:
  - With multiple senders, the amount of information that can be transmitted in equilm depends on fine details such as: the shape of the boundary of the state space, how similar preferences of the senders are,...
  - Also properties of the state space and sender preferences cannot be investigated independently if one allows state-dependent preferences

4.3 Good News and Bad News

- Milgrom (Bell, 1981)
- Monotonicity plays a crucial role in information economics: e.g. Spence-Mirrlees condition/single-crossing, monotonicity of the bidding function in auctions,...
- Milgrom: "...it is surprising that studies of rational expectations equilibria and of the problem of moral hazard make no use of any such property...Such results has, unfortunately, been out of reach because no device has been availabel for modelling 'good news."
- And he provides one
- Moreover, he provides a set of tools which are indispensable in modelling information problems in a *very* wide class (eg. Friedman and Holden (AER, 2007))
- Let $\Theta \subseteq \mathbb{R}$ be the set of possible values of a random parameter $\tilde{\theta}$

- The set of possible signals about $\tilde{\theta}$ is $X \subseteq \mathbb{R}^m$
- Let $f(x|\theta)$ be the conditional density

Definition 13. *A signal x is* ***more favorable*** *than the signal y if for every nondegenerate prior distribution G the posterior distribution $G(\cdot|x)$ first order stochastically dominates $G(\cdot|y)$*

- ie.
$$\int U(\theta)\, dG_1(\theta) > \int U(\theta)\, dG_2(\theta),$$
or equivalently $G_1(\theta) \leq G_2(\theta)$ with at least one inequality strict.
- Suppose G is a prior that assigns probabilities $g(\theta)$ and $g\left(\bar{\theta}\right)$ to θ and $\bar{\theta}$. By Bayes' Rule
$$\frac{g\left(\bar{\theta}|x\right)}{g(\theta|x)} = \frac{g\left(\bar{\theta}\right)}{g(\theta)} \frac{f\left(x|\bar{\theta}\right)}{f(x|\theta)}, \tag{7}$$
and similarly for signal y.

Theorem 8. *x is more favorable than y if and only if for every $\bar{\theta} > \theta$*
$$f\left(x|\bar{\theta}\right) f(y|\theta) - f(x|\theta) f\left(y|\bar{\theta}\right) > 0.$$

Proof. Suppose $g(\theta) = g\left(\bar{\theta}\right) = 1/2$. The by the definition of more favorable and (7) we have
$$\frac{f\left(x|\bar{\theta}\right)}{f(x|\theta)} > \frac{f\left(y|\bar{\theta}\right)}{f(y|\theta)}.$$
Fix G and consider θ^* that $0 < G(\theta^*) < 1$. For $\theta \leq \theta^*$ the above inequality implies
$$\frac{\int_{\bar{\theta}>\theta^*} f\left(x|\bar{\theta}\right) dG\left(\bar{\theta}\right)}{f(x|\theta)} > \frac{\int_{\bar{\theta}>\theta^*} f\left(y|\bar{\theta}\right) dG\left(\bar{\theta}\right)}{f(y|\theta)}$$
Flipping the inequality
$$\frac{f(x|\theta)}{\int_{\bar{\theta}>\theta^*} f\left(x|\bar{\theta}\right) dG\left(\bar{\theta}\right)} < \frac{f(y|\theta)}{\int_{\bar{\theta}>\theta^*} f\left(y|\bar{\theta}\right) dG\left(\bar{\theta}\right)}$$
Integrating over θ for $\theta \leq \theta^*$ yields
$$\frac{\int_{\theta\leq\theta^*} f(x|\theta)\, dG(\theta)}{\int_{\bar{\theta}>\theta^*} f\left(x|\bar{\theta}\right) dG\left(\bar{\theta}\right)} < \frac{\int_{\theta\leq\theta^*} f(y|\theta)\, dG(\theta)}{\int_{\bar{\theta}>\theta^*} f\left(y|\bar{\theta}\right) dG\left(\bar{\theta}\right)},$$
which is equivalent to
$$\frac{G(\theta^*|x)}{1 - G(\theta^*|x)} < \frac{G(\theta^*|y)}{1 - G(\theta^*|y)}$$
which in turn implies that $G(\theta^*|x) < G(\theta^*|y)$ □

Definition 14. *Let $X \subseteq \mathbb{R}$. The family of densitites $\{f(\cdot|\theta)\}$ have the (strict)* ***Monotone Likelihood Ratio Property (MLRP)*** *if for every $x > y$ and $\bar{\theta} > \theta$, $f\left(x|\bar{\theta}\right) f(y|\theta) - f(x|\theta) f\left(y|\bar{\theta}\right) > 0$.*

Remark 12. *This property holds for many families of densities, although it is trivial to construct counterexamples. It holds for, among others the: normal, uniform, exponential, Poisson and chi-squared distributions.*

Theorem 9. *The family of densities $\{f(\cdot|\theta)\}$ has the strict MLRP if and only if $x > y$ implies that x is more favorable than y.*

Proof. Immediate from the definition and the previous theorem. □

Definition 15. *Two signals are* ***equivalent*** *if $f\left(x|\bar{\theta}\right) f(y|\theta) - f(x|\theta) f\left(y|\bar{\theta}\right) = 0$.*

Definition 16. *If two signals are equivalent or one is more favorable than the other then they are* ***comparable.***

- The following theorem establishes the very useful fact that information system which is comparable can be modelled as a real-valued variable with the MLRP

Theorem 10. *Suppose that two signals in X are comparable. Then there exists a function $H : X \to \mathbb{R}$ such that $H(\tilde{x})$ is a sufficient statistic for $\tilde{x}$ and such that the densities of $H(\tilde{x})$ have the strict MLRP.*

Proof. Let $h : \mathbb{R} \to \mathbb{R}$ be a bounded increasing function and let

$$H(x) = \int h(\theta dG(\theta|x)),$$

where G is a prior over θ. By comparability $H(x) > H(y) \Leftrightarrow x$ if more favorable than y. Then by the previous theorem the densities have the strict MLRP. And since $H(x) = H(y) \Leftrightarrow x$ and y are equivalent, $H(\tilde{x})$ is a sufficient statistic. □

- Now consider one application of this from Milgrom (1981)–persuasion games
- An interested party provides information to a decision maker in order to influence her decision
- Suppose there is a B(uyer) and a S(eller)
- Good is of unknown quality $\tilde{\theta}$
- B's payoff is $\tilde{\theta} F(q) - pq$, where q is quantity and p is price
- S's payoff is an arbitrary increasing function of q
- Suppose S has a vector of information about the product $\tilde{x} = (\tilde{x}_1, ..., \tilde{x}_N)$
- These can be concealed, but not misreported
- A report by S is a subset of $\mathbb{R}^N$ which asserts that $\tilde{x} \in \mathbb{R}^N$
- A reporting strategy is a function f from $\mathbb{R}^N$ to subsets of $\mathbb{R}^N$

- A precise report has $r(x) = \{x\}$ and a vague one has $r(x) = \mathbb{R}^N$
- Solution concept: sequential equilibrium
- Full disclosure is a report whereby S does not conceal any information which is payoff relevant to B
- Milgrom shows that in every equilbrium of this game the strategy chosen by S is one of full disclosure
- Key idea: any information which is withheld is interpreted as negative information– why didn't S report it if it wasn't?! (this deep idea also appears in Hart (1980), Grossman (1981)
- What if information is costly to communicate?
- Suppose that B can only listen to $k < N$ observations
- Also suppose that $\tilde{x}_1, ..., \tilde{x}_N$ are conditionally independent and that F has the MLRP
- In this setting S always reports the k most favorable observations in any sequential equilibrium
- Intuition: want to maximize B's expectation of the quality and by MLRP this is done by sending the most favorable messages
- Again, extreme skepticism/assume the worst

5 Moral Hazard

5.1 Introduction

- Many applications of principal-agent problems
  - Owner / Manager
  - Manager / Worker
  - Patient / Doctor
  - Client / Lawyer
  - Customer / Firm
  - Insurer / Insured
- History:
  - Arrow ('60s)
  - Pauly (68), Spence-Zeckhauser
  - Ross (early '70s)
  - Mirrlees (mid '70s)
  - Holmström ('79)
  - Grossman-Hart ('83)

5.2 The Basic Principal-Agent Problem

5.2.1 A Fairly General Model

- $a \in A$ (Action Set)
- This leads to q (verifiable revenue)
- Stochastic relationship $F(q;a)$
- Incentive scheme $I(q)$
- The Principal solves the following problem:

$$\max_{\widehat{I}(\cdot),\widehat{a}} \left\{ \int \left(q - \widehat{I}(q) \right) dF(q;\widehat{a}) \right\}$$

$$s.t.(i)\ \widehat{a} \text{ solves } \max_{a \in A} \left\{ \int u(a, \widehat{I}(q)) dF(q;a) \right\} \tag{ICC}$$

$$(ii)\ \int u(\widehat{a}, \widehat{I}(a)) dF(q;\widehat{a}) \geq \overline{U} \tag{PC}$$

- Use the deterministic problem of the Principal inducing the Agent to choose the action because there may be multiple actions which are equivalent for the Agent but the Principal might prefer one of them
- The Principal is really just a risk-sharing device

5.2.2 The First-Order Approach

- Suppose $A \subseteq \mathbb{R}$
- The problem is now

$$\max_{a,I(\cdot)} \left\{ \int_{\underline{q}}^{\bar{q}} (q - I(q)) f(q|a) dq \right\}$$

$$\text{subject to}$$

$$a \ \in \arg\max_{\hat{a} \in A} \left\{ \int_{\underline{q}}^{\bar{q}} u(I(q)) f(q\,|a|) dq - G(a) \right\} \tag{ICC}$$

$$\int_{\underline{q}}^{\bar{q}} u(I(q)) f(q\,|a|) dq - G(a) > \overline{U} \tag{PC}$$

- IC looks like a tricky object
- Maybe we can just use the FOC of the agent's problem
- That's what Spence-Zeckhauser, Ross, Harris-Raviv did

- FOC is
$$\int_{\underline{q}}^{\bar{q}} u(I(q))f_a(q|a)dq = G'(a)$$
- SOC is
$$\int_{\underline{q}}^{\bar{q}} u(I(q))f_{aa}(q|a)dq = G''(a)$$
- If we use the first-order condition approach:
$$\begin{aligned} \frac{\partial}{\partial I} &= 0 \Rightarrow -f(q;a) + \mu u'(I(q))f_a(q|a) + \lambda u'(I(q))f(q|a)) = 0 \\ &\Rightarrow \frac{1}{u'(I(q))} = \lambda + \mu \frac{f_a(q;a)}{f(q;a)} \end{aligned}$$
- f_a/f is the likelihood ratio
- $I \uparrow q \Leftrightarrow \frac{f_a}{f} \uparrow q$
- But the FOC approach is not always valid – you are throwing away all the global constraints
- The $I(q)$ in the agent's problem is endogenous!
- MLRP $\Rightarrow$ "the higher the income the more likely it was generated by high effort"

Condition 1 (Monotonic Likelihood Ratio Property ("MLRP")). *(Strict) MLRP holds if, given $a, a' \in A$, $a' \preceq a \Rightarrow \pi_i(a')/\pi_i(a)$ is decreasing in i.*

Remark 13. *It is well known that MLRP is a stronger condition than FOSD (in that MLRP $\Rightarrow$ FOSD, but FOSD $\not\Rightarrow$ MLRP).*

Condition 2 (Covexity of the Distribution Function Condition). $F_{aa} \geq 0$.

Remark 14. *This is an awkward and somewhat unnatural condition–and it has little or no economic interpretation. The CDFC holds for no known family of distributions*

- MLRP and CDFC ensure that it will be valid (see Mirrlees 1975, Grossman and Hart 1983, Rogerson 1985)
- FOC approach valid when FOC≡ICC
- In general they will be equivalent when the Agent has a convex problem
- To see why (roughly) they do the trick suppose that $I(q)$ is almost everywhere differentiable (although since it's endogenous there's no good reason to believe that)
  - The agent maximizes
$$\int_{\underline{q}}^{\bar{q}} u(I(q))f(q|a)dq - G(a)$$

- Integrate by parts to obtain

$$u(I(\bar{q})) - \int_{\underline{q}}^{\bar{q}} u'\left(I\left(q\right)\right) I'\left(q\right) F(q|a) dq - G(a)$$

- Now differentiate twice w.r.t. a to obtain

$$-\int_{\underline{q}}^{\bar{q}} u'\left(I\left(q\right)\right) I'\left(q\right) F_{aa}(q|a) dq - G''(a) \qquad (*)$$

- MLRP implies that $I'(q) \geq 0$
- CDFC says that $F_{aa}(q|a) \geq 0$
- $G''(a)$ is convex by assumption
- So (*) is negative

- Jewitt's (Ecta, 1988) assumptions also ensure this by restricting the Agent's utility function such that this is the case
- Grossman and Hart (Ecta, 1983), proposed the LDFC, (initially referred to as the Spanning Condition).
- Mirrlees and Grossman-Hart conditions focus on the Agent controlling a family of distributions and utilize the fact that the ICC is equivalent to the FOC when the family of distributions controlled by the Agent is one-dimensional in the distribution space (which the LDFC ensures), or where the solution is equivalent to a problem with a one-dimensional family (which the CDFC plus MLRP ensure)

Remark 15. *Single-dimensionality in the distribution space is not equivalent to the Agent having a single control variable – because it gets convexified*

- It is easy to see why the LDFC works because it ensures that the integral in the IC constraint is linear in e.

5.2.3 Beyond the First-Order Approach: Grossman-Hart

Grossman-Hart with 2 Actions

- Grossman-Hart (Ecta, 1983)
- Main idea of GH approach: split the problem into two step
  - Step 1: figure out the lowest cost way to implement a given action
  - Step 2: pick the action which maximizes the difference between the benefits and costs
- $A = \{a_L, a_H\}$ where $a_L < a_H$ (in general we use the FB cost to order actions–this induces a complete order over A if A is compact)
- Assume $q = q_1 < ... < q_n$
- Note: a finite number of states

- Payment from principal to agent is I_i in state i
- $a_H \rightarrow (\pi_1(a_H), ..., \pi_n(a_H))$
- $a_L \rightarrow (\pi_1(a_L), ..., \pi_n(a_L))$
- Agent has a v-NM utility function $U(a, I) = V(I) - G(a)$
- $G(a_H) > G(a_L)$
- Reservation utility of $\overline{U}$
- Assume V defined on $(\underline{I}, \infty)$
- $V' > 0, V'' < 0, \lim_{I \to \underline{I}} V(I) = -\infty$ (avoid corner solutions, like $\ln(I)$ instead of $I^{1/2}$)
- Of course, a legitimate v-NM utility function has to be bounded above and below (a result due to Arrow), but...

First Best (a verifiable):

- Define $h \equiv V^{-1}$
- $V(h(V)) = V$
- Pick a
- Let $C_{FB}(a) = h(\overline{U} + G(a))$
- since $V(I) - G(a) = \overline{U}$, $V(I) = G(a) + \overline{U}, I = h(\overline{U} + G(a))$
- Can write the problem as

$$\max_{a \in A} \{ {}_{i=1}^{n} \pi_i(a) q_i - C_{FB}(a) \}$$

Second Best:

- $a = a_L$ then pay you $C_{FB}(a_L)$ regardless of the outcome
- $a = a_H$

$$\min_{I, \ldots, I_n} \left\{ \sum_{i=1}^{n} \pi_i(a_H) I_i \right\}$$

$$s.t. (i) \sum_{i=1}^{n} \pi_i(a_H) V(I_i) - G(a_H) \geq \sum_{i=1}^{n} \pi_i(a_L) V(I_i) - G(a_L) \qquad \text{(ICC)}$$

$$(ii) \sum_{i=1}^{n} \pi_i(a_H) V(I_i) - G(a_H) \geq \overline{U} \qquad \text{(PC)}$$

- We use the Vs as control variables (which is OK since V is strictly increasing in I)

- $v_i = V(I_i)$

$$\min_{v_1,\ldots,v_n} \left\{ \sum_{i=1}^{n} \pi_i(a_H) h(v_i) \right\} \tag{*}$$

$$s.t.(i) \sum_{i=1}^{n} \pi_i(a_H) v_i - G(a_H) \geq \sum_{i=1}^{n} \pi_i(a_L) v_i - G(a_L) \tag{ICC}$$

$$(ii) \sum_{i=1}^{n} \pi_i(a_H) v_i - G(a_H) \geq \overline{U} \tag{PC}$$

- Now this is just a convex programming problem
- Note, however, that the constraint set is unbounded – need to be careful about the existence of a solution

Claim 1. *Assume $\pi_i(a_H) > 0, \forall i$. Then $\exists$ a unique solution to (*)*

Proof. (sketch): The only way there could not be a solution would be if there was an unbounded sequence $(v_1', ..., v_n') \Rightarrow I$s are unbounded above $\Rightarrow VarI \to \infty$, where $I_i = h(v_i)$. $\underset{\sim}{V}$ unbounded $\Rightarrow \underset{\sim}{I}$ unbounded above (if not Is$\to \underline{I}$ and vs$\to -\infty \Rightarrow$PC violated. With $V(\cdot)$ strictly concave $E[\underset{\sim}{I}] \to \infty$ as $\underset{\sim}{I} \to \infty$ if $\underline{I} \neq -\infty$. If $\underline{I} = -\infty$ the PC will be violated because of risk-aversion. □

- Solution must be unique because of strict convexity with linear constraints
- π_is are all positive
- Let the minimized value be $C(a_H)$
- Compare $\sum_{i=1}^{n} \pi_i(a_H) q_i - C(a_H)$ to $\sum_{i=1}^{n} \pi_i(a_L) q_i - C_{FB}(a_L)$
- This determines whether you want a_H or a_L in the second-best

Claim 2. $C(a_H) > C_{FB}(a_H)$ *if* $G(a_H) > G(a_L)$. *The second-best is strictly worse than the first-best if you want them to take the harder action.*

Proof. (sketch): Otherwise the ICC would be violated because all of the π_is are positive and so all the vs would have to be equal - which implies perfect insurance. □

Claim 3. *The PC is binding*

Proof. (sketch): If $\sum_{i=1}^{n} \pi_i(a_H) v_i - G(a_H) > \overline{U}$ then we can reduce all the v_is by ε and the Principal is better off without disrupting the ICC. □

- FB=SB if:

1. Shirking is optimal
2. V is linear and the agent is wealthy (risk neutrality) – make the Agent the residual claimant (but need to avoid the wealth constraint)

3. $\exists i$ sth $\pi_i(a_H) = 0, \pi_i(a_L) > 0$ (MOVING SUPPORT). If the Agent works hard they are perfectly insured, if not they get killed.

- Now form the Lagrangian:

$$\begin{aligned} = \quad & \sum_{i=1}^{n} \pi_i(a_H)h(v_i) \\ & -\mu \left(\sum_{i=1}^{n} \pi_i(a_H)v_i - G(a_H) - \sum_{i=1}^{n} \pi_i(a_L)v_i + G(a_L) \right) \\ & -\lambda \left(\sum_{i=1}^{n} \pi_i(a_H)v_i - G(a_H) \right) \end{aligned}$$

- The FOCs are:

$$\frac{\partial}{\partial v_i} = 0, \forall i$$

$$\pi_i(a_H)h'(v_i) - \mu\pi_i(a_H) + \mu\pi_i(a_L) - \lambda\pi_i(a_H) = 0$$

$$\frac{1}{V(I_i)} = h'(v_i) = \lambda + \mu - \mu\frac{\pi_i(a_L)}{\pi_i(a_H)} \quad \forall i = 1, ..., n$$

- Note that $\mu > 0$ since if it was not then $h'(v_i) = \lambda$ which would imply that the v_is are all the same, thus violating the ICC
- Implication: Payments to the Agent depend on the likelihood ratio $\frac{\pi_i(a_L)}{\pi_i(a_H)}$

Theorem 11. *In the Two Action Case, Necessary and Sufficient conditions for a monotonic incentive scheme is the MLRP*

- This is because the FOC approach is valid the in the 2 action case even w/out the CDFC
- *This behaves like a statistical inference problem even though it is not one (because the actions are endogenous)*
- Linearity would be a very fortuitous outcome
- Note: in equilm the Principal knows exactly how much effort is exerted and the deviations of performance from expectation are stochastic – but this is optimal *ex ante*

5.2.4 Random Schemes

- Can one do better with random schemes? Do you want to add noise ?
- Suppose the Principal decided to "flip a coin", $j \in \{1, ..., m\} \rightarrow pr(j) = q(j)$

- $\pi_{ij}(a) = q_j \pi_i(a)$
- Suppose w_{ij} was the optimal scheme and let $\widetilde{w}_i$ be the certainty equivalent:

$$u(\widetilde{w}_i) = \sum_j q_j u(w_{ij}) \quad, \forall i$$

- But we haven't changed the ICC or PC
- However, the Principal has cost $\widetilde{w}_i$ and $\widetilde{w}_i < \sum_j q_j w_{ij}$ due to the concavity of $u(\cdot)$. So the Principal is better off. Contradiction
- Therefore random schemes cannot be better
- They put more risk onto the risk-averse Agent and that requires the Agent to be compensated for bearing that risk
- Can also use the sufficient statistic result - the random scheme adds no information about the likelihood ratio (and generalizes to the case where the Principal is risk-averse)

5.2.5 Linear Contracts

- Very little that you can say in a general moral hazard model (Grossman and Hart 83)
- Say $w = t + vq$
- Assume normally distributed performance and CARA (exponential) utility
- Let $q = a + \varepsilon$ with $\varepsilon \sim N(0, \sigma^2)$
- Assume the Principal is risk-neutral
- The Agent is risk-averse with:

$$U(w, a) = -e^{-r(w - \psi(a))}$$

- Let $\psi(a) = \frac{ca^2}{2}$
- Note that r is the coefficient of absolute risk-aversion $-u''/u'$
- The Principal solves:

$$\max_{a,t,v} E[q - w]$$
$$s.t. (i) E[-e^{-r(w-\psi(a))}] \geq -e^{-r\overline{w}} \qquad \text{(PC)}$$
$$(ii) a \in \arg\max_a E[-e^{-r(w-\psi(a))}] \qquad \text{(ICC)}$$

- Let $x \sim N(0, \sigma_x^2)$

- $E[e^{\gamma x}] = e^{\gamma^2 \sigma_x^2/2}$ (this is essentially the calculation done to yield the moment generating function of the normal distribution – see Varian for a more detailed derivation)

$$\begin{aligned} & E[-e^{-r(w-\psi(a))}] \\ = & -E[-e^{-r(t+va+v\varepsilon-\psi(a))}] \\ = & -e^{-r(t+va-\psi(a))}E[e^{-rv\varepsilon}] \\ = & e^{-r\widehat{w}(a)} \end{aligned}$$

- $\widehat{w}(a) = t + va - \frac{r}{2}v^2\sigma^2 - \frac{1}{2}ca^2$
- Now $\max_a \{\widehat{w}(a)\}$
- FOC is $v - ca = 0 \Rightarrow a = v/c$
- Replace a with v/c in the Principal's Problem and they solve:

$$\max_{v,t} \left\{ \frac{v}{c} - (t + \frac{v^2}{c}) \right\} \tag{8}$$

$$s.t. \widehat{w}(a) = \widehat{w}(\frac{v}{c}) = \overline{w}PC \tag{9}$$

- The PC is, written more fully:

$$t + \frac{v^2}{c} - \frac{r}{2}v^2\sigma^2 - \frac{v^2}{2c}$$

- ie.

$$t + \frac{v^2}{2c} - \frac{r}{2}v^2\sigma^2 = \overline{w}$$

- Substituting for t:

$$\max_v \left\{ \frac{v}{c} - \frac{v^2}{c} + \frac{v^2}{2c} - \frac{r}{2}v^2\sigma^2 - \overline{w} \right\}$$

- The FOC is:

$$\frac{1}{c} - \frac{v}{c} - rv\sigma^2 = 0$$

- Hence:

$$v = \frac{1}{1 + rc\sigma^2}$$

- Which is a nice, simple, closed form solution
- But the linearity restriction is not at all innocuous
- In fact, linear contracts are not optimal in this setting!
- Without the restriction one may approximate the first-best

Example 1: Moving Support

- $q = a + \varepsilon$ and ε is uniformly distributed on $[-k, k]$ with $k > 0$
- So the Agent's action moves the support of q

Claim 4. *The first-best can be implemented by a non-linear contract*

Proof. Let a^* be the first-best level of effort. q will take values in $[a^* - k, a^* + k]$. Scheme: pay w^* whenever $q \in [a^* - k, a^* + k]$ and pay $-\infty$ otherwise. Just a Mirrlees Scheme (which is certainly not linear) □

- With bounded support the Principal can rule out certain outcomes *provided the Agent chooses the FB action.*

EXAMPLE 2:

- $q = a + \varepsilon$ and $\varepsilon \sim N[0, \sigma^2]$

$$\Rightarrow f(q, a) = \frac{1}{(2\pi\sigma)^{1/2}} e^{-(q-a)^2/2\sigma^2}$$

- Calculate the likelihood ratio:

$$f_a(q, a) = -\frac{1}{(2\pi\sigma)^{1/2}} e^{-(q-a)^2/2\sigma^2} \times \frac{-(q-a)}{\sigma^2}$$

- $\frac{f_a}{f} = \frac{q-a}{\sigma^2}$
- as $q \to \infty^+$, $\frac{f_a}{f} \to \infty$
- So the likelihood ratio can take on values on $(-\infty, \infty)$
- For extreme values (ie. in the tails of the distn) the Principal gets almost perfect information

Claim 5. *FB a^* can be arbitrarily approximated*

Proof. Suppose the Principal chooses an incentive scheme as follows: if $q < \underline{q} \to$ low transfer k, if $q \geq \underline{q} \to$ transfer w^*. Suppose the Agent has a utility function $u(y), u'(y) > 0, u''(y) < 0$ and cost of effort $\psi(a)$. To implement a^* under the above scheme we need that:

$$IC : \int_{-\infty}^{\underline{q}} u(k) f_a(q, a^*) dq + \int_{\underline{q}}^{\infty} u(w^*(q)) f_a(q, a^*) dq = \psi'(a^*)$$

But this violates the PC by:

$$l = \int_{-\infty}^{\underline{q}} \left[u(w^*(q)) - u(k)\right] f(q^*) dq$$

□

Claim 6. *One can choose $\underline{q}$ and k to make l arbitrarily small.*

Proof. Given $-M, \exists \underline{q}$ such that:

$$\frac{f_a(q,a)}{f(q,a)} \leq -M \quad \text{for } q \leq \underline{q}$$

$\Rightarrow \frac{f_a}{f}\left(\frac{-1}{M}\right) \geq 1 \Leftrightarrow f \leq f_a(\frac{-1}{M})$

$$\begin{aligned} \Rightarrow \quad l &\leq \int_{-\infty}^{\underline{q}} \left[u(w^*(q)) - u(k)\right] f_a(q^*, a) \left(\frac{-1}{M}\right) dq \\ &= \frac{-1}{M}(\cdot) \end{aligned}$$

Therefore one can make l arbitrarily small by making M arbitrarily large □

- The *expected* punishment is bounded away from ∞
- Mirrlees's (1974) idea again – this time without the moving support
- Although the size of the punishment grows, its frequency falls at a faster rate

5.3 Multi-Agent Moral Hazard

5.3.1 Relative Performance Evaluation

- Holmström (Bell, 1982)
- Assume for simplicity 2 symmetric agents
- $q_1 = a_1 + \varepsilon_1 + \beta\varepsilon_2$
- $q_2 = a_2 + \varepsilon_2 + \beta\varepsilon_1$
- ε_1 and ε_2 are *iid* $N\left(0, \sigma^2\right)$
- Principal is risk-neutral
- Agents are risk-averse
- Agents have utility functions of the form:

$$U(a,w) = -e^{-r(w-\psi(a))}$$

- where $\psi(a) = \frac{1}{2}ca^2$
- Assume linear contracts so that:

$$w_1 = t_1 + v_1 q_1 + u_1 q_2$$

$$w_2 = t_2 + v_2 q_2 + u_2 q_1$$

- $u_1 = u_2 = 0$ is the case of no relative performance evaluation
- The Principal solves:

$$\max_{a_1,t_1,v_1,u_1} E[q_1 - w_1]$$
$$s.t.(i) \quad E[-e^{-r(w_1-\frac{1}{2}ca^2)}] \geq -e^{-r\overline{w}} \qquad \text{(PC)}$$
$$(ii) \quad a_1 \in \arg\max_a E[-e^{-r(w_1-\frac{1}{2}ca^2)}] \qquad \text{(ICC)}$$

- where $\overline{w}$ is the reservation wage
- The Agent's payoff is $\widetilde{w}_1 - \frac{1}{2}ca_1^2$

$$\begin{aligned} &= t_1 + v_1q_1 + u_1q_2 - \frac{1}{2}ca_1^2 \\ &= t_1 + v_1(a_1 + \varepsilon_1 + \beta\varepsilon_2) + u_1(a_2 + \varepsilon_2 + \beta\varepsilon_1) - \frac{1}{2}ca_1^2 \end{aligned}$$

- The certainty equivalent ("CE") of this is:

$$CE = t_1 + v_1a_1 + u_1a_2 - \frac{r}{2}\sigma^2((v_1 + \beta u_1)^2 + (u_1 + \beta v_1)^2) - \frac{1}{2}ca_1^2$$

since

$$\begin{aligned} risk &= var\left(v_1(a_1 + \varepsilon_1 + \beta\varepsilon_2) + u_1(a_2 + \varepsilon_2 + \beta\varepsilon_1)\right) \\ &= var\left(v_1(\varepsilon_1 + \beta\varepsilon_2) + u_1(\varepsilon_2 + \beta\varepsilon_1)\right) \\ &= \sigma^2\left[(v_1 + \beta u_1)^2 + (u_1 + \beta v_1)^2\right] \end{aligned}$$

- The FOC is:

$$\begin{aligned} \frac{\partial CE}{\partial a_1} &\Rightarrow v_1 = ca_1 \\ a_1 &= \frac{v_1}{c} \end{aligned}$$

- The Principal solves:

$$\max_{t_1,v_1,a_1} \left\{\frac{v_1}{c} - (t_1 + v_1\frac{v_1}{c} + u_1a_2)\right\}$$
$$s.t. CE = \overline{w}$$

- Which is equivalent to:

$$\max_{v_1,u_1} \left\{ \begin{array}{c} \frac{v_1}{c} - \frac{v_1^2}{c} - \overline{w} + \frac{v_1^2}{c} - u_1a_2 + u_1a_2 \\ -\frac{r}{2}\sigma^2((v_1 + \beta u_1)^2 + (u_1 + \beta v_1)^2) - \frac{1}{2}c\frac{v_1^2}{c^2} \end{array} \right\}$$

- Simplification yields:

$$\max_{v_1,u_1}\left\{\frac{v_1}{c}-\frac{v_1^2}{2c}-\frac{r}{2}\sigma^2((v_1+\beta u_1)^2+(u_1+\beta v_1)^2)\right\}$$

- Trick: given v_1, u_1 is determined to minimize risk; then v_1 is set to trade-off risk sharing and incentives
- Fix v_1 and solve:

$$\min_{u_1}\left\{(v_1+\beta u_1)^2+(u_1+\beta v_1)^2\right\}$$

$$\Rightarrow 2(v_1+\beta u_1)+2(u_1+\beta v_1)=0$$

$$u_1=\frac{-2\beta}{1+\beta^2}v_1$$

- $u_1 = 0$ if $\beta = 0$ (ie. the environments are completely independent)
- Filter out the common shock
- This implies that:

$$v_1=\frac{1}{1+r\sigma^2 c\frac{(1-\beta^2)^2}{1+\beta^2}}$$

- It doesn't matter whether $\beta \lesseqgtr 0$
- You can make incentives more high powered because there is a way to insure the Agent
- Noise is netted out

5.3.2 Moral Hazard in Teams

- Holmström (Bell, 1982)
- n agents $1, ..., n$ who choose actions $a_1, ..., a_n$
- This produces revenue $q(a_1, ..., a_n)$ with $q(\cdot)$ concave
- Agent's utility function is $I_i - \psi_i(a_i)$ with $\psi_i(\cdot)$ convex
- In the first-best:

$$\max\left\{q(a_1,...,a_n)-\sum_{i=1}^{n}\psi_i(a_i)\right\}$$

- The FOC is:

$$\frac{\partial q}{\partial a_i}=\psi_i'(a_i) \quad ,\forall i$$

- In the second-best assume that a_i is observable only to agent i but that q is observable to everyone

- A partnership consists of sharing rules $s_i(a_i), i = 1, ..., n$ such that

$$\sum_i s_i(q) \equiv q \tag{10}$$

- Might suppose that $s_i(q) \geq 0, \forall i$
- In a Nash Equilibrium each agent solves:

$$\max_{a_i} \{s_i(q(a_i, a_{-i})) - \psi_i(a_i)\}$$

- The FOC is:

$$s_i'(q)\frac{\partial q(a_i, a_{-i})}{\partial a_i} = \psi_i'(a_i)$$

- Need $s_i'(q) = 1, \forall i$ to get the FB
- But we know from (10) that $\sum_i s_i'(q) \equiv 1$
- Can't get the FB
- Nothing to do with risk-aversion – there is no uncertainty here
- Say we introduce an $(n+1)$th party such that:

$$s_i(q) \equiv q(a^*) - F_i, \quad \forall i = 1, ..., n$$

$$s_{n+1}(q) = \sum_i F_i - nq(a^*)$$

- This will be profitable for the $(n+1)$th party if we pick F_i such that $\sum_{i=1}^n F_i + q(a^*) \geq nq(a^*)$
- And also profitable for the agents if $F_i \leq q(a^*) - \psi_i(a_i^*)$
- These can both be satisfied because at the FB $q(a^*) - \sum_{i=1}^n \psi_i(a_i^*) > 0$
- We have made everyone the residual claimant
- However, the $(n+1)$th party wants it to fail. They might burn the factory down, ... Call them the Budget Breaker ("BB")
- They might also collude with one of the Agents
- A side contract between BB and i – this merges BB and i into one person and we are back into the n agent case
- n people could collude to "borrow" q and game the BB
- This mechanism (making everyone the residual claimant) is similar to Groves-Clarke we we saw earlier

5.3.3 Random Schemes

- Legros & Matthews (Restud, 1993)
- Can get the FB by using a random scheme
- Say $n = 2$, $A_i \in \{0, 2\}$, $q(a) = a_1 + a_2, \psi_i = \frac{a_i^2}{2}$
- FB:

$$\max_{a_1, a_2} \left\{ a_1 + a_2 - \frac{1}{2}a_1^2 - \frac{1}{2}a_2^2 \right\}$$

- $\Rightarrow a_1^* = a_2^* = 1$
- SB: Agent 1 $a_1^* = 1$, Agent 2 $a_2^* = 1$ with $\Pr(1 - 2\delta)$, 2 wth $\Pr(\delta)$, 0 wtih $\Pr(\delta)$

$$s_2(q) = \frac{(q-1)^2}{2}$$

$$s_1(q) = q - \frac{(q-1)^2}{2}$$

- Agent 2 indifferent about a_2 given $a_1^* = 1$
- Her payoff is:

$$\frac{(1 + a_2 - 1)^2}{2} - \frac{1}{2}a_2^2 = 0$$

- So Agent 2 is perfectly prepared to play the mixed strategy
- Make 1 pay 2 a large fine if $q < 1$ or $q > 3$
- Wealth constraint is a very serious problem
- Also give 2 a big incentive not to play the mixed strategy – and as always, it is very hard to make sure that mixed strategies are verifiable
- It is in general a big problem – as n goes large effort $\to 0$ since FOC$\Rightarrow \frac{1}{n} = \psi'(a)$

5.3.4 Tournaments

- Lazear and Rosen (JPE, 1981)
- Let $q_1 = a_1 + \varepsilon_1$, $q_2 = a_2 + \varepsilon_2$
- Assume that $\varepsilon_1, \varepsilon_2$ are $iid \sim N(0, \sigma^2)$
- The Principal and Agent are risk-neutral
- The winner of the tournament is the one with the higher q and gets the prize w – and both get a fixed payment t

- Agent i solves:

$$\begin{aligned}\max_{a_i}\{E[w_i-\psi(a_i)]\} &= \max_{a_i}\{t+pw-\psi(a_i)\}\\ &= \max_{a_i}\{t+pr(q_i>q_j)w-\psi(a_i)\}\\ &= \max_{a_i}\{t+pr(a_i+\varepsilon_i>a_j+\varepsilon_j)w-\psi(a_i)\}\\ &= \max_{a_i}\{t+pr(\varepsilon_j-\varepsilon_i<a_i-a_j)w-\psi(a_i)\}\\ &= \max_{a_i}\{t+G(a_i-a_j)w-\psi(a_i)\}\end{aligned}$$

- where G is the CDF of $\varepsilon_j-\varepsilon_i$
- Note that $\varepsilon_j-\varepsilon_i\sim N(0,2\sigma^2)$
- FOCs:

$$g(a_i-a_j)w-\psi'(a_i)=0$$

  and:

$$g(a_j-a_i)w-\psi'(a_j)=0$$

- The symmetric Nash Equilibrium is $a_i=a_j=a$
- Hence $g(0)w=\psi'(a)$
- FB $\Rightarrow\psi'(a^{FB})=1$
- Therefore:

$$w=\frac{1}{g(0)}$$

- *Just need an ordinal measure to get the FB*
- Risk-neutrality seems like a huge issue – with both risk-neutral we could just use residual claimancy anyway
- With risk-aversion and a common shock consider a comparison of the piece-rate and the tournament
  - With a big common shock (ie. $(\varepsilon_1,\varepsilon_2)\sim N(0,\Sigma)$, where $\Sigma=\begin{pmatrix}\sigma^2 & R\\ R & \sigma^2\end{pmatrix}$ the tournament dominates because the piece-rate doesn't take into account the common shock
  - With a small common shock the tournament imposes lots of risk and is thus dominated by the piece-rate scheme
  - See Green & Stokey (1983)
- Green-Stokey setup
  - P pays a prize w_i to the individual who places ith in the tournament
  - $\pi=\sum_{i=1}^{n}(q_i-w_i)$

- Assume that individuals are homogeneous in ability
- If individual j exerts effort e_j, her output is given by $q_j = e_j + \varepsilon_j + \eta$, where ε_j and η are random variables with mean zero and distributed according to distributions F and G respectively
- Assume that F and G are statistically independent
- Refer to η as the "common shock" to output and ε_j as the "idiosyncratic shock" to output
- Each agent's tility is given by: $u(w_i) - c(e_j)$ where $u' \geq 0, u'' \leq 0, c' \geq 0, c'' \geq 0$
- Time 0: the principal commits to a prize schedule $\{w_i\}_{i=1}^n$. Time 1: individuals decide whether or not to participate. Time 2: if everyone has agreed to participate at time 1, individuals choose how much effort to exert. Time 3: output is realized and prizes are awarded
- Restrict attention to symmetric pure strategy equilibria
- A unique symmetric pure strategy equilibrium will always exist, provided that the distribution of idiosyncratic shocks is "sufficiently dispersed"
- In a symmetric equilibrium, every individual will exert effort e^*
- Furthermore, every individual has an equal chance of winning any prize and the expected utility is

$$\frac{1}{n} \sum_i u(w_i) - c(e^*)$$

- In order for it to be worthwhile for an individual to participate in the tournament, it is necessary that

$$\frac{1}{n} \sum_i u(w_i) - c(e^*) \geq \bar{U}$$

- An individual who exerts effort e while everyone else exerts effort e^* receives expected utility

$$\begin{aligned} U(e, e^*) &= \sum_i \varphi_i(e, e^*) u(w_i) - c(e) \\ \text{where } \varphi_i(e, e^*) &= \Pr(i\text{th place}|e, e^*), \end{aligned}$$

- that is, the probability of achieving place i given effort e while all other agents choose effort e^*
- Each agent chooses e to maximize $U(e, e^*)$
- The first-order condition for this problem is

$$c'(e) = \sum_i \frac{\partial}{\partial e} \varphi_i(e, e^*) u(w_i)$$

- Since we know that the solution to the maximization problem is $e = e^*$, it follows that

$$c'(e^*) = \sum_i \beta_i u(w_i)$$
$$\text{where } \beta_i = \frac{\partial}{\partial e}\varphi_i(e, e^*)\Big|_{e=e^*}$$

– Note that β_i does not depend upon e^* but simply upon the distribution function F

– Routine results from the study of order-statistics imply that the formula for β_i is

$$\beta_i = \binom{n-1}{i-1} \int_{\mathbb{R}} F(x)^{n-i-1}(1-F(x))^{i-2}\left((n-i)-(n-1)F(x)\right)f(x)^2 dx$$

Since $\sum_i \varphi_i = 1$, it follows that $\sum_i \beta_i = 0$

– It is also easily shown that if F is a symmetric distribution ($F(-x) = 1 - F(x)$), that $\beta_i = -\beta_{n+1-i}$.

– Now that we have elaborated the agent's problem, we turn to the principal's problem. The principal's object is to maximize

$$E(\pi) = \sum_j e_j - \sum_i w_i = n\left(e^* - \frac{1}{n}\sum_i w_i\right).$$

– The problem of the principal can therefore be stated as follows

$$\max_{w_i}\left(e^* - \frac{1}{n}\sum_i w_i\right)$$

subject to

$$\frac{1}{n}\sum_i u(w_i) - c(e^*) \geq \bar{U} \qquad \text{(IR)}$$

$$c'(e^*) = \sum_i \beta_i u(w_i) \qquad \text{(IC)}$$

- Tournaments generally suboptimal because they throw away the cardinal information
- RPE individual contracts do better
- Green-Stokey limit result: as the number of players goes to infinite the tournament goes to the optimal RPE individual contract

5.3.5 Supervision & Collusion

- Tirole (JLEO, 1986)
- Consider a Principal who wants a service from an Agent
- Suppose that the supply cost c is 0 or 1 with equal probability

- The value of the service to the Principal is s
- The Agent knows c, the Principal does not
- Assume that the Principal has all the bargaining power
- Assume $s > Z$ so that $s - 1 > s/2$ and hence makes a take-it-or-leave-it offer of 1
- Suppose that the Principal can hire a supervisor at cost z
- The supervisor, with probability p, gets hard evidence that $c = 0$ when $c = 0$
- Assume that the evidence can be destroyed, but that it cannot be falsified / fabricated

Case I: Honest Supervisor

- Optimal contract is: if Supervisor reports $c = 0$ the Agent gets 0, if not the Agent gets 1
- The Principal's payoff is $\frac{1}{2}ps + (1 - \frac{p}{2})(s - 1) - z$
- This is greater than $s - 1$ if z is small enough

Case II: Corrupt Supervisor

- Collusion technology: The Agent can pay the Supervisor t but the Supervisor only receives kt where $k \in [0, 1]$ - but other than this, the side relationship is binding
- Tirole introduces the *Collusion Proofness Principle* – a bit like the Revelation Principle
- Optimal Contract: if produce hard evidence then the Principal pays w
- $w \geq k$ to avoid collusion, because there is 1 on the table for the Agent which is worth k to the supervisor. Obviously $w = k$
- The Principal's payoff, assuming $z = 0$, is:

$$\frac{p}{2}(s - k) + (1 - \frac{p}{2})(s - 1)$$

- With an honest supervisor the payoff is:

$$\frac{ps}{2} + (1 - \frac{p}{2})(s - 1)$$

- The Principal's payoff without a supervisor is $s - 1$
- Since $k < 1$ you always want a supervisor

Comments:

1. Collusion proof principle is not that robust – for instance, if k is random
2. Collusion v. costly effort – rotation of supervisors might be good (make k go down)
3. In some cases you could make the supervisor the residual claimant (eg. speeding fine and police)

5.3.6 Hierarchies

- Qian (Restud, 1994)
- Consider a hierarchy which consists of T layers with top layer being tier 0 and bottom T
- Define the "span of control" as s_{i+1} which is the number of individuals in tier $i+1$ who are monitored by an individual in tier i
- Let the number of individuals in a tier be X_i
- Assume that $X_T = N$ and that output is given by $\theta N y_T$ where θ is a measure of profitability, N is the scale and y_T is the effective output per final worker
- Assume that there is a "loss of control" represented by $y_t = a_t y_{t-1}$ with $a_t \leq 1$
- Assume that the Principal has no incentive problem so that $a_0 = 1$ and that for all other tiers there is a convex effort cost $g(a) = a^3$
- Let w be the wage and $p = 1/s_i$ be the probability of getting caught when shirking
- The program for the optimal organizational design is:

$$\max_{s_t,a_t,x_t,T} \left\{ \theta N y_t - \sum_{t=1}^{T} g(a_t) s_t x_t \right\}$$
$$s.t.(i) \quad x_t = x_{t-1} s_t$$
$$(ii) \; y_t - y_{t-1} a_t$$
$$(iii) \; x_0 = y_0 = 1, x_T = N$$
$$(iv) \; 0 \leq a_t \leq 1, \forall t$$

Results:

Proposition 7. *Assume that $T = 1$ which means that there is one Principal and N workers so that $y_1 = y_0 a = a, x_1 = x_0 s = N, s_1 = N$. Then $a = \min\left\{1, (\theta/3)^{1/2} N^{-1/2}\right\}$.*

Proof. Introducing some new notation, the program to be solved is now:

$$\Pi_1 = \max_{a \in [0,1]} \left\{\theta N a - a^3 N^2\right\}$$

□

The first-order condition is:

$$\begin{aligned} \theta N &= 3a^2 N^2 \\ &\Rightarrow \quad a = \min\left\{1, (\theta/3)^{1/2} N^{-1/2}\right\} \end{aligned}$$

Proposition 8. *Now assume that $T = 2$. Then $a_1^* = 1$ and $a_2^* \leq a_1^*$.*

Proof. Note that $T = 2 \Rightarrow N = x_2 = x_1 s_2$, $x_1 = x_0 s_1 = s_1, y_1 = y_0 a_1 = a_1, y_2 = y_1 a_2 = a_1 a_2$. Now write an unconstrained program with a_1, a_2, x_1 as control variables as follows:

$$\max_{a_1,a_2,x_1} \left\{\theta N a_1 a_2 - a_1^3 x_1^2 - a_2^3 N^2/x_1\right\} \tag{11}$$

First fix a_1 and a_2 and optimize with respect to x_1. This yields the first-order condition:

$$-2x_1 a_1^3 + a_2^3 N^2/x_1^2 = 0$$

This implies:

$$\begin{aligned} x_1 &= \frac{a_2^3 N^2}{2a_1^3} \\ &= \frac{a_2}{a_1}\left(\frac{N^2}{2}\right)^{1/3} \end{aligned}$$

Now substitute into (11):

$$\max_{a_1,a_2} \left\{\theta N a_1 a_2 - a_1 a_2^2 \left[\left(\frac{N^2}{2}\right)^{2/3} + \left(\frac{2}{N^2}\right)^{1/3} N^2\right]\right\}$$

Note that:

$$\left(\frac{N^2}{2}\right)^{2/3} + \left(\frac{2}{N^2}\right)^{1/3} N^2 = N^{4/3}\left(2^{-2/3} + 2^{1/3}\right)$$

and denote $\left(2^{-2/3} + 2^{1/3}\right)$ as $z < 2$. This directly implies that $a_1^* = 1$ since if it were less than 1 then increasing it to 1 and reducing a_2 to keep $a_1 a_2$ constant increases the maximand. Therefore $a_2^* \leq a_1^* = 1$. □

- This means that effort goes down when one moves down the hierarchy. Intuitively, the higher up the hierarchy, the more y_T's are being affect by effort which raises the marginal benefit of effort as one moves up the hierarchy.
- Now we can offer a necessary and sufficient condition for profit under $T = 2$ to be greater than under $T = 1$.

$$\Pi_2 = \max_{a_2} \left\{\theta N a_2 - a_2^2 N^{4/3} z\right\}$$

- Solving for a_2 from above yields:

$$a_2 = \min\left\{1, (\theta/2z)\, N^{-1/3}\right\}$$

and then:

$$\Pi_2 > \Pi_1 \Leftrightarrow \theta^{1/2} N^{1/6} \geq 3$$

- This means that it is optimal to increase the number of layers in the hierarchy if N becomes sufficiently large. The intuition for this is as follows. An increase in N means less supervision for $T = 1$, and therefore a reduction in effort since compensating the decrease in supervision by an increase in wages is prohibitively costly. So an increase in N increases the gain of having an intermediate layer so as to save on wages at the bottom of the hierarchy.

Proposition 9. *The amount of wage inequality w_1/w_2 is increasing in N.*

Proof. With $T = 1$ we have:

$$\begin{aligned} w &= g(a)s \\ &= g(a)N \\ &= \left(\frac{\theta}{3}\right)^{3/2} N^{-1/2} \end{aligned}$$

With $T = 2$:

$$\begin{aligned} w_1 &= g(a_1)s_1 \\ &= a_1^3 \frac{a_2}{a_1} \left(\frac{N^2}{2}\right)^{1/3} \\ &= a_2 \left(\frac{N^2}{2}\right)^{1/3} \\ &= \left(\frac{\theta}{2z}\right) N^{1/3} 2^{1/3} \end{aligned}$$

and:

$$\begin{aligned} w_2 &= g(a_2)s_2 \\ &= g(a_2)\frac{N}{x_1} \\ &= \left(\frac{\theta}{2z}\right)^2 N^{-1/3} 2^{1/3} \end{aligned}$$

It is therefore clear that $\frac{\partial w_1}{\partial N} > 0$ and $\frac{\partial w_2}{\partial N} < 0$. This directly implies that $\partial \left(w_1/w_2\right)/\partial N > 0$. Therefore wage inequality is increasing in N. □

5.4 Moral Hazard with Multiple Tasks

5.4.1 Holmström-Milgrom

- Holmström-Milgrom (JLEO, 1991)
- Different tasks with different degrees of measurability
- Suppose the Agent can sell the Principal's product or someone else's product
- 2 tasks $i = 1, 2$
- Let $q_i = a_i + \varepsilon_i$

- $(\varepsilon_1, \varepsilon_2) \sim N(0, \Sigma)$ where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & R \\ R & \sigma_2^2 \end{pmatrix}$$

- Let the Agent's utility be given by:

$$-e^{-r(w-\psi(a_1,a_2))}$$

- where $\psi(a_1, a_2) = \frac{1}{2}(c_1 a_1^2 + c_2 a_2^2) + \delta a_1 a_2$
- if $\delta > 0$ then the two tasks are technological substitutes, if $\delta < 0$ they are complements
- Assume a linear incentive scheme:

$$w = t + v_1 q_1 + v_2 q_2$$

$$\begin{aligned}
\widehat{w}(a_1, a_2) &= E[w(a_1, a_2)] - \frac{r}{2} var\left(w(a_1, a_2)\right) - \psi(a_1, a_2) \\
&= E[t + v_1\left(a_1 + \varepsilon_1\right) + v_2\left(a_2 + \varepsilon_2\right)] \\
&\quad - \frac{r}{2} var(t + v_1\left(a_1 + \varepsilon_1\right) + v_2\left(a_2 + \varepsilon_2\right)) \\
&\quad - \frac{1}{2}((c_1 a_1^2 + c_2 a_2^2) + \delta a_1 a_2)
\end{aligned}$$

- $E[t + v_1\left(a_1 + \varepsilon_1\right) + v_2\left(a_2 + \varepsilon_2\right)] = t + v_1 q_1 + v_2 q_2$
- $Var(\cdot) = v_1^2 \sigma_1^2 + v_2^2 \sigma_2^2 + 2R v_1 v_2$
- The Agent solves:

$$\max_{a_1, a_2} \{\widehat{w}(a_1, a_2)\}$$

- Let $R = 0$
- The FOCs are now:

$$v_1 = c_1 a_1 + \delta a_2$$

$$v_2 = c_2 a_2 + \delta a_1$$

- Using the FOC approach the Principal solves:

$$\max_{v_1, v_2, a_1, a_2} \{E[q - w] = a_1 + a_2 - t - v_1 a_1 - v_2 a_2\}$$

$$\begin{gathered}
s.t.(i)\ \widehat{w}(a_1, a_2) = t + v_1 a_1 + v_2 a_2 \\
-\frac{r}{2}\left(v_1^2 \sigma_1^2 + v_2^2 \sigma_2^2 + 2R v_1 v_2\right) \geq \overline{W} \\
(ii)\quad v_1 = c_1 a_1 + \delta a_2 \\
(iii)\quad v_2 = c_2 a_2 + \delta a_1
\end{gathered}$$

- (i) must bind so we have:

$$\max_{v_1,v_2,a_1,a_2}\left\{\begin{array}{c} a_1+a_2-\frac{r}{2}\left(v_1^2\sigma_1^2+v_2^2\sigma_2^2+2Rv_1v_2\right) \\ -\frac{1}{2}\left(c_1a_1^2+c_2a_2^2\right)-\delta a_1a_2 \end{array}\right\}$$

$$s.t.\ v_1=c_1a_1+\delta a_2$$

$$v_2=c_2a_2+\delta a_1$$

- FOC1:

$$1-r\sigma_1^2v_1c_1-r\sigma_2^2v_2\delta-v_1=0$$

- $\Rightarrow$

$$v_1=\frac{1-r\sigma_2^2v_2\delta}{1+r\sigma_1^2v_1c_1}$$

$$v_2=\frac{1-r\sigma_1^2v_1\delta}{1+r\sigma_2^2v_2c_2}$$

- Solving simultaneously yields:

$$v_1=\frac{1+r\sigma_2^2(c_2-\delta)}{1+r\sigma_1^2c_1+r\sigma_2^2c_2+r^2\sigma_1^2\sigma_2^2(c_1c_2-\delta^2)}$$

- and symmetrically for v_2

Results:

1. Go from $\delta=1$ to $\delta=-1$ (ie. substitutes to compliments) and v_1,v_2 increase
2. When $\delta=0$:

$$v_1=\frac{1}{1+r\sigma_1^2c_1}$$

   which is simply the one-task case.

3. As $\sigma_2^2\to\infty$ (task 2 is really hard to measure) then:

$$v_2\to 0$$

$$v_1\to\frac{r(c_2-\delta)}{rc_2+r^2\sigma_1^2(c_1c_2-\delta^2)}$$

   Put all the incentive on task 1.

5.5 Dynamic Moral Hazard

5.5.1 Stationarity and Linearity of Contracts

With repetition of the Principal-Agent Problem:

1. Agent may become less risk-averse because of self-insurance (saving)
2. Principal gets more observations to infer effort – less noise
3. Agent has more actions - intertemporal substitution of effort is possible

Holmström and Milgrom (Econometrica, 1987):

- Continuous time
- Wiener Process:
$$dx(t) = \mu(t)dt + \sigma dB(t)$$
- $t \in [0, 1]$
- $x(t)$ = total revenue up to time t
- B is a standard Brownian Motion $x(1) \sim N(\mu, \sigma^2)$
- Principal is risk-neutral
- Agent has CARA utility given by:
$$-e^{-r(s - \int_0^1 c(\mu(t)dt))}$$
- Assumes that the Agent is not saving
- Only the Agent sees the path $[x(t)]_0^1$ - if not then the optimal scheme would be a Mirrlees scheme
- We will consider a two period version of the model - can be generalized (HM 87)
- Three dates $\{0, 1, 2\}$
- Between dates 0 and 1 (period 1) action a_1 is taken
- Between dates 1 and 2 (period 2) action a_2 is taken
- At date 2 the Agent is paid s
- $a_1 \to x = x_1, ..., x_n$ with probabilities $\pi_1(a_1), ..., \pi_n(a_1)$
- $a_2 \to x = x_1, ..., x_n$ with probabilities $\pi_1(a_2), ..., \pi_n(a_2)$
- where x is revenue
- The shocks are independent
- The Agent's utility is $-e^{-r(s-a_1-a_2)}$
- The Principal's payoff is $x_i + x_j$

- The Principal Solves:

$$\max_{\{a_1, \widehat{a}(x_i), s_{ij}\}} \left\{ \sum_i \sum_j \pi_i(\widehat{a}_i)\pi_j(\widehat{a}(x_i))(x_i + x_j - s_{ij}) \right\}$$

$$s.t.(i) \quad \widehat{a}_1(\widehat{a}(x_1)) \in \arg\max \left\{ \sum_i \pi_i(a_i) \sum_j \pi_j(a(x_i)) \left(-e^{-r(s_{ij} - a(x_i) - a_1)} \right) \right\} \qquad (*)$$

$$(ii) \quad \sum_i \pi_i(a_i) \sum_j \pi_j(a(x_i)) \left(-e^{-r(s_{ij} - a(x_i) - a_1)} \right) \geq \overline{U}$$

- Now write the part of (*) which is required to be in the argmax as:

$$\sum_i \pi_i(a_i) e^{ra_1} \left[\sum_j \pi_j(a(x_i)) \left(-e^{-r(s_{ij} - a(x_i))} \right) \right]$$

- Now replace the IC constraint with:

$$\widehat{a}(x_i) \in \arg\max \sum_j \pi_j(a(x_i)) \left(-e^{-r(s_{ij} - a(x_i))} \right)$$

$$\widehat{a}_1 \in \arg\max \sum_i \pi_i(a_i) \sum_j \pi_j(a(x_i)) \left(-e^{-r(s_{ij} - a(x_i) - a_1)} \right)$$

- Call the solution to the overall problem

$$\left(a_1^*, (a^*(x_i)), s_{ij}^* \right)$$

- Let $\overline{U}_i = \sum_j \pi_j(a^*(x_i)) \left(-e^{-r(s_{ij} - a^*(x_i)} \right)$

Claim 7. $\forall i \ \left(a_1^*, (a^*(x_i)), s_{ij}^* \right)$ *must solve:*

$$\max_{\widehat{a}, (\widehat{s}_{ij})_j} \left\{ \sum_j \pi_j \left(\widehat{a} \right) \left(x_j - \widehat{s}_{ij} \right) \right\}$$

$$s.t.(i) \quad \widehat{a} \in \arg\max \left\{ \sum_j \pi_j \left(a \right) \left(-e^{-r(\widehat{s}_{ij} - a)} \right) \right\}$$

$$(ii) \quad \sum_j \pi_j \left(\widehat{a} \right) \left(-e^{-r(\widehat{s}_{ij} - \widehat{a})} \right) \geq \overline{U}_i$$

Proof. (Sketch): Suppose not. Replace $\left(a^* \left((x_i) \right), s_{ij}^* \right)$ with $\left(\overline{a} \left((x_i) \right), \overline{s}_{ij} \right)$ which contradicts the fact that $\left(a_1^*, (a^*(x_i)), s_{ij}^* \right)$ is an optimum. □

- Now, $a^*(x_i) \equiv a^*$
- $s^*_{ij} = s^*_j + k_i$
- *The second period action does not depend on the first period action*
- Now we can write:

$$\max \left\{ \sum_i \pi_i(\widehat{a}) \left(x_i + \sum_j \pi_j(a^*)(x_j - s^*_j - k_i) \right) \right\}$$

$$s.t.(i) \quad \widehat{a} \in \arg\max \left\{ \sum_i \pi_i(a) e^{ra_1} \sum_j \pi_j(a^*) \left(-e^{-r(s^*_i + i - a^*)} \right) \right\} \tag{12}$$

$$(ii) \quad \sum_i \pi_i(\widehat{a}) \left(-e^{-r(k_i - \widehat{a}} \right) \geq \overline{V} \tag{13}$$

- Noting that (12) reduces to $\sum_i \pi_i(a) - e^{-r(k_i - \widehat{a})}$
- The problem above is really just the one period problem again
- So: $\widehat{a} = a^*$
- $k_i = s^*_i + \alpha$
- *The actions should be the same in both periods*
- $s^*_{ij} = s^*_i + \alpha + s^*_j = \left(s^*_i + \frac{\alpha}{2}\right) + \left(s^*_j + \frac{\alpha}{2}\right)$
- It is as if the incentive scheme in each period is:

$$\left(s^*_i + \frac{\alpha}{2}, ..., s^*_n + \frac{\alpha}{2} \right)$$

- Note how difficult it was to get the problem to be stationary

Observations:

1. Keep n accounts: final payment is a linear function of the accounts – the order does not matter. *s is linear in the accounts, but not in x, even though x is linear in the accounts. Up to this point we only have a constant scheme, not a linear one.* Note that s is not a sufficient statistic for x.

$$s^* = N_1 w^*_1 + N_2 w^*_2 + ...N_n w^*_n + \alpha$$

   - But $N_1 = Q$

2. *To get linearity one needs just two possible outcomes per period*
3. The sufficient statistic result doesn't hold here – we're throwing away some information here

4. As $t \to \infty$ we don't converge to the first-best

- Now extend to the continuous time Brownian Motion case
- Brownian Motion where the action affects the mean, not the variance
- Can be approximated by a two outcome, discrete process, repeated a lot of times (because the Central Limit Theorem says that one can approximate by binomials)
- *The optimal scheme will be linear in the limit because of the two outcome per period result*
- But strong assumptions: (i) control the mean not the variance, (ii) CARA, (iii) No consumption until the end
- Folds back into static schemes and multi-tasking–justifies linear contracts (although dynamic v. statics settings!)
- Hellwig and Schmidt discrete time approximation

5.5.2 Renegotiation

- Return to a basic Principal-Agent setup
- Suppose there are two outputs $q_1 < q_2$ and two actions $a_L < a_H$
- Let $p(a) = \Pr(q_2|a) \to p(a_j) = p_j$, where $j = L, H$
- So $\Pr(q_2|a_H) = p_H$ and $\Pr(q_2|a_L) = p_L$
- Cost of effort given by: $\psi(a_H) = K > \psi(a_L) = 0$
- Suppose further that there is a lag between action choice and the outcome

Case I:

- Action not observed by the principal (Fudenberg & Tirole (Ecta, 1990))
- Expect renegotiation because the action is sunk (should have the principal buy-out the risky position and improve risk-sharing)
  - But this might have bad incentive effects ex ante
- Key observation: to avoid full insurance *ex post*, it must be that the principal remains unsure about which action the agent chose
- The optimal contract must induce randomization by the agent so that there remains asymmetric information at the bargaining stage
- The timeline of the game is as follows:
  - t=0: Contract on $\{w_1(\widehat{a}), w_2(\widehat{a})\}$

- t=1: Agent chooses $a = \begin{cases} a_H \text{ w/ } pr\ x \\ a_L \text{ w/ } pr\ 1-x \end{cases}$
- t=2: Principal renegotiates and offers $\{\widehat{w}_1(\widehat{a}), \widehat{w}_2(\widehat{a})\}$, where $\widehat{a}$ is announced effort
- t=3: Output is realized and payments made

- Suppose the principal wants to implement a_H
- Suppose the incentive scheme was $I = \alpha + \beta q$, with $\beta > 0$
- Say the Agent chose a_H and suppose that the principal has all the bargaining power
- $\alpha + \beta q \to \widehat{\alpha}$
- But knowing that they will get $\widehat{\alpha}$, they will choose a_L
- Fudenberg and Tirole show that you can sustain a mixed strategy equilibrium – there's asymmetric information in the bargaining stage which provides some incentive to work / put in some effort
- wlog can restrict attention to renegotiation-proof contracts (ie. such that there does not exist another contract which also satisfies the PC and ICC and makes the principal strictly better-off
  - Suppose P offered a contract C' which was not RP, this contract would be replaced by C'' which is RP and since the agent anticipates renegotiation her choice of x is unchanged
- Usual screening logic – ICC binding for $a = a_L$ agents (because they want to pretend to be $a = a_H$), not for $a = a_H$ agents
- So $w_1(a_L) = w_2(a_L) = w^* \Rightarrow$ full insurance for a_L agents
- And if $x > 0$ is optimal then $w_1(a_H) < w_2(a_H)$
- Furthermore: $u(w^*) = p_L u(w_2(a_H)) + (1 - p_L)u(w_1(a_H))$
- This stems from the fact that the interim ICC for the low type is binding. That constraint is:

$$p_L u(w_2(a_L)) + (1 - p_L) u(w_1(a_L)) \geq p_L u(w_2(a_H)) + (1 - p_L) u(w_1(a_H))$$

- So we have

$$u(w^*) = p_L u(w_2(a_H)) + (1 - p_L)u(w_1(a_H)) \tag{ICC-L}$$

- At the initial stage the contract is $C = (x = 1, (w_j(a)))$ is not RP – if it was then it would induce P to offer full insurance to type H agents $\Rightarrow w_2(a_H) = w_1(a_H) = w(a_H)$, but then by ICC_L we have $u(w^*) = u(w(a_H)) \Rightarrow w^* > w(a_H) \Rightarrow$ ex ante ICC violated in contradiction of $x = 1$
- In fact, given w^*, there is a maximum value of x (ie. $x(w^*)$) that can be induced by a RP contract

- Ex ante ICC:
$$p_H u(w_2(a_H)) + (1-p_H)u(w_1(a_H)) - K = u(w^*)$$

- Note that if $x \neq 0, 1$ then the agent is indifferent b/w $x = 0$ and $x = 1$ because she anticipates no renegotiation and P, expecting A to choose the stipulated x will not renegotiate the contract. Therefore the ex ante ICC is binding

- ICC and the ex ante IC constraint jointly determine $w_2(a_H)$ and $w_1(a_H)$ as a fn of $w^* \Rightarrow w_2(w^*), w_1(w^*)$

- Then P chooses w^* to maximize expected profits subject to PC and ICC:

$$\max_{w^*} \left\{ \begin{array}{c} x(w^*)[p_H(q_2 - w_2(w^*)) + (1-p_H)(q_1 - w_1(w^*))] \\ +(1-x(w^*))[p_L q_2 + (1-p_L)q_1 - w^*] \end{array} \right\}$$
$$s.t. (i) PC, (ii) ICC, (iii) RP$$

- Suppose P increases w^* by dw^* small - then she can provide better insurance to type H without violating the ex post ICC

- So $w_2 -> w_2 + dw_2$ with $dw_2 < 0$ and $w_1 -> w_1 + dw_1$ with $dw_1 > 0$, where $p_H dw_2 u'(w_2(w^*)) + (1-p_H) dw_1 u'(w_1(w^*)) = 0$ (just subtract $IC(w+dw)$ from $IC(w)$)

- But the initial contract is RP, so P is indifferent at the margin which implies the following:
$$x(w^*)[p_H dw_2 + (1-p_H)dw_1] = (1-x(w^*))dw^* \qquad \text{(RP)}$$

- And if we know the functions $w_1(w^*)$ and $w_2(w^*)$ then we can find $x(w^*)$

Case II:

- Action is observed by the principal after it is taken but before the resolution of uncertainty

- Now the renegotiation is good (Hermalin & Katz (Ecta, 1991))

- In fact you can achieve the FB

- The principal offers a fixed wage which depends on the observed effort level

- Suppose the agent has all the bargaining power

- Set $I(q) = q - F$

- eg. $F = 0$ (if they also have all the bargaining power ex ante)

- Agent chooses a, principal sees this then the agent sells the random output to the principal

- $q \to W$

- Perfect insurance / risk-sharing and perfect incentives

5.6 Career Concerns

5.6.1 Overview

- Formal incentive schemes are not the only way of motivating people
- Takeovers, debt, product market competition, implicit contracts, labor market competition (ie. career concerns)
- Work hard – get a good reputation
- Fama (JPE, 1980): sort of claimed that CCs would lead to the first-best – a bit extreme

5.6.2 Holmstrom's Model

- Formal analysis developed by Holmstrom (Essays in Honor of Lars Wahlbeck '82, then reprinted in Restud in '99)
- 2 period version (the general case is quite impressive)
- Risk-neutral principal ("Employer") and a risk-neutral Agent ("Manager")
- $y_t = \theta + a_t + \varepsilon_t$
- $t \in \{1, 2\}$
- θ_t is the manager's ability
- a_t is her action
- ε_t is white noise
- Symmetric information other than effort observation (only M sees that) – in particular, M doesn't know her own ability so that contracting takes places under symmetric information
- $\theta \sim N(\bar{\theta}, \sigma_\theta^2)$
- $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$
- $\theta, \varepsilon_1, \varepsilon_2$ are independent
- M can move costlessly at the end of the period and there is a competitive market for M's services (same technology)
- Cost of effort $\psi(a), \psi'(a) > 0, \psi''(a)$ – and assume that $\psi(0) = 0$ and that $\psi'(0) = 0$
- Discount factor δ
- Market observes y_1 and y_2 but they are not verifiable – so can't contract on them
- Can only pay a fixed wage in each period
- With a one period model the reputation effect is absent – no incentive to work at all $\rightarrow$ get a flat wage and set $a_1 = 0 \Rightarrow y_1 = \theta + \varepsilon_1$

- Therefore $E[y_1] = E[\theta] = \bar{\theta}$
- Since there is perfect competition $w = \bar{\theta}$
- Take w_2 to be set by competition for M's services and note that $a_2 = 0$ because it is the last period

$$\begin{aligned} w_2 &= E[y_2 \mid \text{info}] \\ &= E[\theta \mid \text{info}] \\ &= E[\theta \mid y_1 = \theta + a_1 + \varepsilon_1] \end{aligned}$$

- Assume that the market has rational expectations about a_1
- Let a_1^* be the equilibrium value of a_1 (a Rational Expectations Equilibrium "REE")

$$\begin{aligned} w_2 &= E[\theta \mid \theta + a_1^* + \varepsilon] \\ &= y_1 - a_1^* \end{aligned}$$

- Update the prior such that:

$$E[\theta|(y_1 - a_1^*)] = \bar{\theta}\left(\frac{\sigma_\varepsilon^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}\right) + (y_1 - a_1^*)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}\right)$$

- Note the effect of the signal to noise ratio
- The first period problem for the Agent is:

$$\max_{a_1} \{w_1 + \delta E[w_2] - \psi(a_1)\}$$

- Which can be written as:

$$\max_{a_1} \left\{ w_1 + \delta\left(\bar{\theta}\left(\frac{\sigma_\varepsilon^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}\right) + \left(\bar{\theta} + a_1 - a_1^*\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}\right)\right) - \psi(a_1) \right\}$$

$$\max_{a_1} \left\{ \delta\left(a_1 - a_1^*\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}\right) - \psi(a_1) \right\}$$

- The FOC is:

$$\delta\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}\right) = \psi'(a_1) \tag{14}$$

- Increasing effort translates into an increased inference of agent talent
- In the FB $\psi'(a_1^{FB}) = 1$
- From (14) we know that $\psi'(a_1) < 1$ because of two things: (i) $\delta < 1$ and (ii) $\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\varepsilon^2}\right) < 1$
- $\Rightarrow 0 < a_1^* < a_1^{FB}$

- The fact that even when the agent does nothing they are valuable in the second period prevents there being a backward induction unraveling – *but relies crucially on the additive technology*

1. $a_1^* \uparrow$ if σ_θ^2 high or σ_ε^2 low
2. Suppose that there are more periods: zero in the last period $\Rightarrow a_t \rightarrow 0$ and $t \rightarrow \infty$
3. Could also (as Holmström does) have ability getting shocked over time – need this to keep the agent working and get out of the problem in 2, above. In equilibrium the market knows how hard M is working – *disciplined with respect to the out of equilibrium moves, but no fooling in equilibrium*
4. Career concerns don't always help you - eg. in multi-tasking model the competitive labor market distorts the relative allocation of time
5. Gibbons & Murphy: looked at CEO incentive schemes - found more formal schemes later in career - empirical confirmation
6. People may work too hard early on: let $y_t = a_t + \theta + \varepsilon_t, t \in \{1, 2, 3\}, \varepsilon_1 \equiv 0, var(\varepsilon_2) > 0, var(\varepsilon_3) > 0$. The FOC for period 1 is $a_2 = a_3 = 0, \delta + \delta^2 = \psi'(a_1)$. The market learns about θ at the end of period 1. $\delta + \delta^2 > 1$ unless δ is smallish

5.6.3 Career Concerns with Multiple Tasks

- Consider an additive normal model as follows:

$$\begin{aligned} y_i &= \theta_i + a_i + \varepsilon_i \\ \theta_i &\sim N(\bar{\theta}, \sigma_\theta^2) \\ \varepsilon_i &\sim N(0, \sigma_\varepsilon^2) \end{aligned}$$

- $i \in \{1, 2\}$
- Talents may be correlated, but the εs are *iid*
- Assume that the market cares about $\theta_1 + \theta_2$
- Define $\widehat{a} = a_1 + a_2$
- $(\theta_1 + \theta_2) \sim N(2\bar{\theta}, 2(1+\rho)\sigma_\theta^2)$ where ρ is the correlation coefficient between θ_1 and θ_2
- Note that $(\varepsilon_1 + \varepsilon_2) \sim N(0, 2\sigma_\varepsilon^2)$
- If the total cost of effort is $\psi(a)$ then we obtain the following FOC:

$$\psi'\left(a^{SB}\right) = \delta \frac{2(1+\rho)\sigma_\theta^2}{2(1+\rho)\sigma_\theta^2 + 2\sigma_\varepsilon^2}$$

- Note that a^{SB} increases with ρ (since an increase in ρ means that there is a higher signal to noise ratio because there is higher initial uncertainty about talent relative to pure noise)

- Implication for cluster of tasks among agents: one agent should be allocated a subset of tasks that require similar talents
- This is very different than under explicit incentives, where you increase effort by reducing uncertainty on talents and therefore uncluster tasks

5.6.4 Instrumental Career Concerns

- Imagine you are an advisor who has some information which is valuable to a decision maker
- Also imagine that you (A) and the decision maker (DM) have the same preferences, so no bias.
- A has an incentive to truthfully reveal her information.
- There would be an issue if DM thinks that you might be biased? How: if there is another type of A who doesn't have the same preferences as DM
- "Good" A doesn't like the idea of being thought to be biased.
- Note that this is not just shoved into the utility function! It is because it affects the degree to which her advice is listened to.
- Now you have an incentive to lie, for repetitional reasons
- So how much information ends up getting conveyed?
- Can be the case if if A is sufficiently concerned about her reputation, the no information is conveyed in equilibrium.
- Key paper here is Morris (JPE, 2001)–"Political Correctness"
- Two periods, 1 and 2.
- In period 1, DM's optimal decision depends on the state of the world which is either 0 or 1 (i.e. $\omega_1 \in \{0, 1\}$).
- Each state occurs with probability 1/2.
- This is DM's prior (and A's for that matter).
- A observes a signal about the true state, formally $s_1 \in \{0, 1\}$.)
- The signal is informative: with probability γ, the signal is equal to the true state (assume that $1/2 < \gamma < 1$.)
- With probability λ, A is "good" in the sense that she has the same preferences as DM. But with probability $1 - \lambda$, A is "bad," in the sense that meaning that she is biased and just always wants DM to take the same decision (not matter what signal A got).
- After seeing the signal A can send a message $m_1 \in \{0, 1\}$.
- DM observes the message and then take an action $a_1 \in \mathbb{R}$.

- DM then sees the true state (and so does A)
- So DM will rationally update about the type of the advisor.
- If A told her the actual state then it is more likely that she is the good type.
- And the fact that A cares about how she is perceived is the whole mechanism which is interesting here.
- Period 2 is just like period 1, the state is ω_2, the message is m_2, and the signal is s_2.
- Signals are *iid* across periods.
- DM's preferences in period i are represented by the payoff function $V_{DM}^i = -(a_i - \omega_i)^2$, so she wants to "hit the state"
- To weight these across periods let the total payoff be $V_{DM} = -x_1(a_1 - \omega_1)^2 - x_2(a_2 - \omega_2)^2$.
- Good A has exactly these preferences.
- Bad A always wants action 1 to be taken: so that $V_{A,Bad} = y_1 a_1 + y_2 a_2$.
- Proceed by backward induction.
- Period 2: there is no more reputation building to be done, so A just focuses on her current goal–just a cheap talk game.
- There is a unique informative equilibrium in the second period of the game.
- Of course, there is also an uninformative one: a "babbling equilibrium" (Crawford and Sobel, 1982).
- Focus on the informative equilm: suppose DM learns something from the message she receives and chooses a higher action after (say) message 1
- Then bad A will want to announce 1 no matter what signal she got.
- Good A will want to truthfully report he signal, Because DM will choose a higher action if she hears 1 than if she announces 0–and this is better for DM (and hence good A).
- Can compute the optimal action of DM in period 2 (given the message and updating about the type of A), and then the value function for each type of A.
- Now work back to period 1.
- Again there is a babbling equilibrium.
- Main focus: do there exist equilibria in which good A truthfully reports.
- Suppose there is such an equilm. Does bad A want to tell the truth (ie. pool).
- No updating about types in such an equilm–no reputation building.

- But bad A wants to convince DM to take action 1, and if there's no reputational cost then bad A will lie to do that. Contradiction.
- Can show that bad A always announces 1 is she gets signal 1, and with probability v if she observes signal 0.
- And good A tells truth.
- So prob. that good A announces 1 when the true state is 1 is γ (pr. that she observed the accurate signal).
- Pr. bad A says 1 when $\omega = 1$ is $\gamma + (1-\gamma)v$, since she gets the high signal with pr γ, and with pr. $1-\gamma$ bad A observes 0 and announces 1 with pr v.
- So, Bayesian updating means that the probability that A is good, after getting message 1 and seeing observed first period state 1 is:

$$\Lambda(\lambda, 1, 1) = \frac{\lambda\gamma}{\lambda\gamma + (1-\lambda)(\gamma + (1-\lambda)v)}. \tag{15}$$

- Since $v > 0$ this probability is less than λ! ie. even though good A told the truth and was right, her reputation goes down!
- It can be shown that

$$\Lambda(\lambda, 0, 1) = \Lambda(\lambda, 0, 0) > \Lambda(\lambda, 1, 1) > \Lambda(\lambda, 1, 0). \tag{16}$$

- Both types of A have a reputational incentive to announce 0 independent of their signal (in fact, even if their signal was perfect!)
- This is true in any informative equilm.
- Good A tells the truth if the signal is zero. If the signal is 1 then there exists $\overline{x_1}$ such that for all $x_1 < \overline{x_1}$ she lies.
- Note that this leads to loss of information, but can't bias the decision–the expected value of DM's ex post belief is his ex ante belief (by the definition of conditional probability).
- If the second period is important enough (x_2 is large enough for a fixed x_1) then no information is conveyed in period 1.
- Can more repetition help? Can we get away from this conclusion that equilm is uninformative precisely when that period doesn't matter much?

5.7 The Value of Information in Agency Problems

5.7.1 Motivating Questions

1. How valuable is better information in an agency relationship?
2. What do we mean by "better" information?

3. Are there differences between the classic model, the multi-task model, and the career concerns model?

4. Can it be good for the principal to add noise?

- Other motivation: there are a number of very useful tools that are used to answer these questions, and it's useful as either a consumer or producer of applied theory to be familiar with them.

5.7.2 Information in the Linear Model

- Recall the Holmstrom-Milgrom linear model
- Say $w = t + vq$
- Assume normally distributed performance and CARA (exponential) utility
- Let $q = a + \varepsilon$ with $\varepsilon \sim N(0, \sigma^2)$
- Assume the Principal is risk-neutral
- The Agent is risk-averse with:

$$U(w, a) = -e^{-r(w-\psi(a))}$$

- Let $\psi(a) = \frac{ca^2}{2}$
- Note that r is the coefficient of absolute risk-aversion $-u''/u'$
- The Principal solves:

$$\max_{a,t,v} E[q - w]$$
$$s.t.(i) E[-e^{-r(w-\psi(a))}] \geq -e^{-r\overline{w}} \quad \text{(IR)}$$
$$(ii) a \in \arg\max_{a} E[-e^{-r(w-\psi(a))}] \quad \text{(IC)}$$

- Let $x \sim N(0, \sigma_x^2)$
- $E[e^{\gamma x}] = e^{\gamma^2 \sigma_x^2 / 2}$ (this is essentially the calculation done to yield the moment generating function of the normal distribution – see Varian for a more detailed derivation)

$$\begin{aligned} & E[-e^{-r(w-\psi(a))}] \\ = & -E[-e^{-r(t+va+v\varepsilon-\psi(a))}] \\ = & -e^{-r(t+va-\psi(a))} E[e^{-rv\varepsilon}] \\ = & e^{-r\widehat{w}(a)} \end{aligned}$$

- $\widehat{w}(a) = t + va - \frac{r}{2} v^2 \sigma^2 - \frac{1}{2} ca^2$
- Now $\max_{a} \{\widehat{w}(a)\}$

- FOC is $v - ca = 0 \Rightarrow a = v/c$
- Replace a with v/c in the Principal's Problem and they solve:

$$\max_{v,t} \left\{ \frac{v}{c} - (t + \frac{v^2}{c}) \right\} \quad (17)$$

$$s.t. \widehat{w}(a) = \widehat{w}(\frac{v}{c}) = \overline{w} \quad (18)$$

- The IR constraint is, written more fully:

$$t + \frac{v^2}{c} - \frac{r}{2} v^2 \sigma^2 - \frac{v^2}{2c}$$

- ie.

$$t + \frac{v^2}{2c} - \frac{r}{2} v^2 \sigma^2 = \overline{w}$$

- Substituting for t:

$$\max_{v} \left\{ \frac{v}{c} - \frac{v^2}{c} + \frac{v^2}{2c} - \frac{r}{2} v^2 \sigma^2 - \overline{w} \right\}$$

- The FOC is:

$$\frac{1}{c} - \frac{v}{c} - rv\sigma^2 = 0$$

- Hence:

$$v = \frac{1}{1 + rc\sigma^2}$$

- As the amount of noise increases, the intensity of incentives goes down, and so does the principal's payoff.
- This conclusion is extremely general in single agent principal-agent models–even when the first-order approach isn't valid (Holden, 2006).
  - Grossman-Hart (1983) show that a Blackwell garbling increases agency costs
  - Kim (1995) shows the same thing for a mean preserving spread of the likelihood ratio (when the first-order approach is valid)

5.7.3 The Sufficient Statistic Theorem

- Recall the definition of a sufficient statistic

Definition 17. *A statistic $T(x)$ is* ***sufficient*** *for a parameter θ if the conditional distribution of the data X, given the statistic $T(x)$, does not depend on the parameter θ. i.e. $pr(X = x|T(x) = t, \theta) = pr(X = x|T(x) = t)$.*

- Setup
  - Agent takes an action $a \in A$, where A is some (possibly high dimensional) compact set.

- Possible outcomes (revenues to P) $\{q_1, ..., q_n\}$.
- Action induces a probability distribution over outcomes, so that the probabilities of the n states are $\pi_1(a), ..., \pi_n(a)$.
- Agent's cost of effort is $\psi(a)$.
- Payment from P to agent in state i is w_i.
- Agent's utility function is u and P's payoff function is V.

- Say there is a signal which is realized after effort is chosen by the Agent but before the realization of the outcome such that :

$$\pi_{ij}(a) = \pi(i, j \mid a)$$

- ie. probability of outcome i, signal j conditional on action a
- Signal does not enter directly into objective functions – only though the probabilities
- Now, letting $\psi(a)$ be the cost of effort, the Principal solves:

$$\max \left\{ \sum_{i,j} \pi_{ij}(a) V(q_i - w_{ij}) \right\}$$

$$s.t.(i) \sum_{i,j} \pi_{ij}(a) u(w_{ij}) - \psi(a) \geq \overline{U} \tag{IR}$$

$$(ii) a \in \arg\max \left\{ \sum_{i,j} \pi_{ij}(a) u(w_{ij}) - \psi(a) \right\} \tag{IC}$$

- Put the Lagrange multiplier λ on the IR
- The IC FOC is $\sum \pi'_{ij}(a) u(w_{ij}) = 1$
- Forming the Lagrangian and finding $\frac{\partial L}{\partial w_i} = 0, \forall i, \forall j$ yields:

$$\frac{V'(q_i - w_{ij})}{u'(w_{ij})} = \lambda + \mu \frac{\pi'_{ij}(a)}{\pi_{ij}(a)} \tag{19}$$

- When is the optimal w_{ij} independent of j ?
- Same as before if

$$\frac{\pi'_{ij}(a)}{\pi_{ij}(a)} = \frac{\pi'_i(a)}{\pi_i(a)}$$

- In the continuum case (denoting the additional signal s this is just:

$$\frac{g_a(q, s|a)}{g(q, s|a)} = \frac{f_a(q|a)}{f(q|a)}$$

- Integrating this object with respect to a means that it is equivalent to the existence of two functions $m(q|a)$ and $n(q|s)$ such that:

$$g(q, s|a) = m(q|a)\, n(q|s).$$

- That is, that q is a *sufficient statistic* for the pair (q, s) with respect to a
- This representation is known as the Halmos-Savage factorization criterion (or theorem) – see DeGroot (1971) for further details
- So, the optimal incentive scheme is conditioned on s if and only if s is informative about a, given that q is already available

5.7.4 Random Schemes

- Can one do better with random schemes? Do you want to add noise ?
- Suppose the Principal decided to "flip a coin", $j \in \{1, ..., m\} \to pr(j) = q(j)$
- $\pi_{ij}(a) = q_j \pi_i(a)$
- Suppose w_i was the optimal scheme and let $\widetilde{w}_i$ be the certainty equivalent:

$$u(\widetilde{w}_i) = \sum_j q_j u(w_{ij}) \quad , \forall i$$

- But we haven't changed the IC or IR
- However, the Principal has cost $\widetilde{w}_i$ and $\widetilde{w}_i < \sum_j q_j w_{ij}$ due to the concavity of $u(\cdot)$. So the Principal is better off. Contradiction
- Therefore random schemes cannot be better
- They put more risk onto the risk-averse Agent and that requires the Agent to be compensated for bearing that risk
- Can also use the sufficient statistic result - the random scheme adds no information about the likelihood ratio (and generalizes to the case where the Principal is risk-averse)

5.7.5 Information in Career Concerns Settings

Setup

- Key paper is Dewatripont-Jewitt-Tirole (Restud, 99a)
- Going to write down a very general version of the career concerns model
- E(mployee) and M(arket)
- E chooses a vector of actions $(a_1, ..., a_n) \in \mathbb{R}^n$ at cost $c(a)$.

- M observes a vector of outputs $(y_1, ..., y_n) \in \mathbb{R}^n$.
- E's payoff is $t - c(a)$.
- t is M's expectation of E's *talent*, represented by the scalar θ.
- The joint density of performance and talent is $f(\theta, y|a)$.
- The marginal density is $\hat{f}(y|a) = \int f(\theta, y|a)d\theta$.
- Given equilibrium action a^* the reward to E is

$$t = E(\theta|y, a^*) = \int \theta \frac{f(\theta, y|a^*)}{\hat{f}(y|a^*)} \tag{20}$$

- The gradient vectors wrt a are c_a and $\hat{f}_a$.
- Holmstrom 82 is the special case where $t = t_2$, $a = a_1$, $y = y_1$, $c(a) = \psi(a)/\delta$ and $f(\theta, y|a)$ is proportional to $exp[-(\theta - \theta^2)/2\sigma_\theta^2]exp[-(y - \theta - a)^2/2\sigma_\epsilon^2]$.

5.7.6 Equilibrium

- E solves $\max_a \{E_y[E_\theta(\theta|y, a^*)] - c(a)\}$.
- The FOC is

$$\frac{d}{da}\left(\int \left(\int \theta \frac{f(\theta, y|a^*)}{f(y|a^*)} d\theta\right) \hat{f}(y|a)dy\right)|_{a=a^*} = c_a(a^*). \tag{21}$$

- This is

$$\int\int \theta f(\theta, y|a^*) \frac{\hat{f}_a(y|a^*)}{\hat{f}(y|a^*)} dy d\theta = c_a(a^*). \tag{22}$$

- Now observe that the likelihood ratio is mean zero, ie. $E(\hat{f}_a/\hat{f}) = 0$.
- So the equilm condition is just

$$cov\left(\theta, \frac{\hat{f}_a}{\hat{f}}\right) = c_a(a^*) \tag{23}$$

- If there is one task, noise is normal and $y = \theta + a + \epsilon$.
- So y is normal with mean $\theta + a$ and variance $\sigma_\theta^2 + \sigma_\epsilon^2$.
- So $\hat{f}(y|a)$ is proportional to

$$\exp\left[-\frac{(y - \overline{\theta} - a)^2}{2(\sigma_\theta^2 + \sigma_\epsilon^2)}\right]. \tag{24}$$

- And the distributional assumptions imply

$$\frac{\hat{f}_a}{\hat{f}} = \frac{(\theta - \overline{\theta}) + \epsilon}{\sigma_\theta^2 + \sigma_\epsilon^2}. \tag{25}$$

- So we get the result from Holmstrom 82

$$cov\left(\theta, \frac{\hat{f}_a}{\hat{f}}\right) = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2}. \tag{26}$$

- The equilm condition implies

$$\psi'^{*}) = \delta \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2}. \tag{27}$$

5.7.7 Information Orders

- The outputs are *experiments* in the language of statistics
- They provide information about a.
- We'll consider three different orders
  - Blackwell
  - Lehman
  - Fisher
- They differ in how strong they are and what distributions they can order
- One may be more appropriate than another in different problems
- If the statistic y with marginal distn $\hat{f}(y|a)$ is *Blackwell More Informative* than the statistic z with marginal distn $\hat{g}(y|a)$ then there exists a joint distribution for (y, z) such that y is sufficient for (y, z).
- This implies the following *martingale condition* for likelihood ratios

$$\frac{\hat{g}_a(z|a)}{\hat{g}(z|a)} = E\left[\frac{\hat{f}_a(y|a)}{\hat{f}(y|a)}|z, a\right]. \tag{28}$$

- The Blackwell ordering doesn't order all distributions
- Eg. it seems natural to think that a uniform distn F with narrower support than another distn G is more informative
- Yet the Blackwell order doesn't rank these two distns
- It's a partial order
- Suppose $y = a + \epsilon$ and $z = a + \eta$.

- Note also: only a normally distributed y can be Blackwell More Informative than a normally distributed z no matter how spread out (or not) the distns are!
- y and z are what Lehman (1988) calls *location experiments*.
- Recall that the Monotone Likelihood Ratio Property (MLRP) states that $\hat{f}_a/\hat{f}$ is increasing in y (output).
- Roughly says that higher and higher actions make higher outputs more and more likely
- Note: stronger than First Order Stochastic Dominance (FOSD)–implies it, but converse not true
- Suppose $\hat{f}(y|a)$ and $\hat{g}(y|a)$ satisfy MLRP
- Then the martingale condition is equivalent to the existence of a function $\mathscr{L}(.,.)$ that is increasing in each argument such that y is equal in distribution to it.
- There are other representations of this (see Jewitt 1997 for details)
- If we are more restrictive and restrict attention to Normal or Exponential distributions then we can use the Fisher criterion
- We say that y is *Fisher more informative* than z if the expectation of the square of the likelihood ratio of g is weakly less than the expectation of the square of the likelihood ratio of f.

5.7.8 Sufficient Statistics and Career Concerns

- We can now ask an interesting question: does Holmstrom's sufficient statistic theorem apply in the career concerns model?

Definition 18. *The statistic y is **talent sufficient** for (y,z) for expected talent if $E(\theta|y,z,a) = E(\theta|y,a)$.*

Theorem 12. *If y is talent sufficient for (y,z), then making z available when y already is does not alter a^*.*

- Just like in the classic model, the agent's (employees) reward is determined by the likelihood ratio.
- Notice that this is not an if and only if result: just a sufficient condition
- But Holmstrom's sufficient statistic theorem *is* an IFF result
- So there are statistics that are irrelevant in career concerns models that are NOT irrelevant in the classic principal agent model.
- This opens the door to the (intriguing) possibility that noise may be good in career concerns settings
- Consider an example of reporting grades
  - Could report the actual grade, or just pass/fail

- Let the grade be $y = \theta + a$, and suppose that $a \in \{0, 1, 5\}$.
- The student doesn't know her talent (nor does the market).
- Flat prior (each talent is equally likely).
- Assume that $a \in \{0, 1\}$ (working versus shirking).
- Let the cost of effort be such that $c(1) - c(0) \in (1/3, 2/3)$.
- With full disclosure the student improves her payoff by working only when $\theta = 0$.
- So the expected gain is 1/3 and hence the student shirks.
- Now suppose that a pass is $y \geq 1$.
- The ex post payoff from shirking is $\frac{2}{3}[(1 + 5)/2] = 2$.
- But from working it is 3, so shirking isn't an equilm.
- Now consider working.
- By working the student gets 2, but by shirking she get 4/3.

- This is just an example, and it is certainly not true that a Blackwell garbling always increases effort in career concerns settings
- To answer the question of when noise helps and when it hurts we need a notion of what it means for two observables to be *similarly ordered*.
- Consider two observables y and z.
- One statistic is $x = (y, z)$ (the full statistic), another is $T(x) = y$.
- Intuitively that given y is available, a higher value of z increases the expectation of θ.
- If MLRP holds then higher effort tends to increase z.
- So having that information available is going to increase effort.
- Conversely, if the expectation of talent is increasing in a statistic that is decreased by effort then having that information available will decrease effort.
- This notion of a "similar ordering" can be made precise using MLRP and the more complicated concept of **affiliation**. Note: this concept is widely used in auction theory, but we won't go into it here. See DWT p193 for the precise defn.
- One can then show that if effort and news are similarly ordered for z, conditional on y, effort is higher when the market has (y, z) rather than just y, and if they are oppositely ordered then effort is higher when just y is available rather than (y, z).

5.8 Relational Contracts

5.8.1 Overview

- Even when some performance measures can be contracted on, subjective assessments of performance play an important role.
- FX traders generate fairly precise measures of performance (their books are marked to market at the end of each trading day), yet subjective bonuses are a significant part of their compensation–e.g. how well do they meet client demands, timing, etc.

- Lincoln Electric a famous example (most used HBS case ever): use a piece rate, but still about half of a typical worker's compensation is a subjective performance bonus.
- Professional services firms: put into one of several "buckets" based on subjective assessments.
- GE (under Jack Welch) formalized this with an: A,B,C,D system.
- Can think of lots of "every day" examples: promotions, compensation, etc.
- Not only within firms; also between firms
  - Alliances
  - Joint ventures
  - Supplier relationships
- Formal contracts can only be written ex ante about terms that can be verified ex post by a third party.
- A *relational contract* is a self-enforcing "contract".
- An idea which has been considered by economists and non-economists (eg. Macaulay, 1963 American Sociological Review; Klein-Leffler, 1981 JPE)
- Use the theory of repeated games and variations of the Folk Theorem to think about what relationships (relational contracts) are self-enforcing.
- Can be conditioned on non-verifiable outcomes and outcomes that may not be possible to be specified ex ante.
- Virtue: ability to adapt to new information as it becomes available, and ability to expand the range of "contractible" variables.
- Vice: (well, fragility) relies on subgame perfection
  - IR constraints need to be satisfied
  - More subtle: how much faith in rationality?
- Recall the finitely repeated prisoners' dilemma: unravels.
- Also recall the infinitely repeated prisoners' dilemma: if the discount factor is sufficiently high then (C,C) can be sustained as a SPE.
- Same principle at work here: if the future (i.e. relationship) is sufficiently valuable then the short-run temptation to cheat can be disciplined.

5.8.2 Observable Effort

- Start with the case where the agent's effort is observable to both P and A.
- This is basically Bull (1987) and Baker-Gibbons-Murphy (1994).
- Easy to think of this as P and A observing a noisy signal and basing a relational contract on that.
- Get's (much) trickier if want to base a relational contract based on a subjective assessment made by one player, but imperfectly observed by the other (see game theory literature on public versus private monitoring–e.g. Kandori-Matushima (ECMA, 1988)).
- Suppose one P(rincipal) and one A(gent).
- Each period A takes and action a which induces a probability distn over outcomes y.
- Let $y \in \{L, H\}$.
- Suppose $a \in [0, 1]$ and $\Pr(y = H|a) = a$.
- y cannot be contracted on (maybe it's too complex to be described to a third party).
- Suppose that the compensation contract has a base salary s and a discretionary bonus b that P promises to pay A if $y = H$.
- Timing
  - P offers A a contract $w = (s, b)$
  - A accepts, or rejects and gets outside option $\bar{w}$
  - If A accepts then she chooses action $a \in [0, 1]$ at private cost $c(a)$; (note that P does NOT observe a)
  - P and A observe y (in fact, it becomes common knowledge).
  - If $y = H$ then P chooses whether or not to pay b.
- P gets $y - w$
- P's discount rate is r
- A's discount rate doesn't matter, because it is only P who is trying to build a reputation
- A's payoff is $w - c(a)$, with $c(a)$ convex and satisfying $\lim_{a \to 1} \{c(a)\} = \infty$.
- In the FB we have

$$a^{FB} \in \arg\max_a \{L + a(H - L) - c(a)\}.$$

- Thus $c'\left(a^{FB}\right) = H - L$.
- If there was just one period then P won't pay a bonus (there's no point building a reputation) and so A will put in zero effort, and so $y = L$.

- If $L < \bar{w}$ then P, anticipating A's effort choice) will not pay $s > 0$, so A will reject the RC.
- Now consider the infinitely repeated game with discount rate r.
- Suppose both P and A play trigger strategies (cooperate if all previous play was cooperate, and if there was ever a defection then defect forever).
- For a given bonus b, then conditional on A believing that P will cooperate (honor the relational contract) A solves

$$\max_a \{s + ab - c(a)\}.$$

- The FOC is $c'(a^*) = b$.
- A's IR constraint is

$$s + a^*(b)\,b - c(a^*(b)) \geq \bar{w}.$$

- P rationally offers the minimum s that A will accept so that P's expected per-period payoff is

$$V(b) = L + a^*(b)(H - L) - c(a^*(b)) - \bar{w}.$$

- The key is whether A thinks P will honor the RC: if $y - H$ will P pay the discretionary bonus b?
- Fix A's strategy
  - If P does not pay b then she gets $H - s$ this period, and zero thereafter
  - If P does pay b then she gets $H - s - b$ this period, but gets the expected profit of the relationship thereafter
- So P pays the bonus IFF the FV of expected profit starting next period is bigger than the bonus, i.e.:

$$(H - s - b) + \frac{1}{r}V(b) \geq (H - s) + \frac{1}{r} \cdot 0,$$

- This is just

$$b \leq \frac{V(b)}{r}. \tag{29}$$

- Thus, the optimal RC chooses b to maximize $V(b)$ subject to the constraint (29)–the reneging constraint.
- When r is low (P is patient/the interest rate is low) we can get the FB by setting $b = H - L$.
- For high interest rates P is not willing to pay any bonus, and so the relationship cannot be sustained.
- For intermediate interest rates we can't get the FB, but the future relationship is valuable enough to sustain some bonus and hence some effort from A.

5.8.3 Aside: Non-Conguent Performance Measures

- Hard to think of m(any) situation(s) where everything that the principal cares about is contractible–in the sense that it will be enforced by a court.
- Kerr (1975):"the folly of rewarding A, while hoping for B."
- A principal who failed to realize the non-congruence of what she can contract on and what she cares about would be making that error.
- Business history is littered with examples of principals who have committed this error.
  - Baker, Gibbons and Murphy (1994) point to H.J. Heinz Company rewarding division managers for increases in earnings over previous year earnings, which lead to manipulation of accounting variables through timing of booking costs and revenues.
  - At Dun & Bradstreet salespeople only received commissions if customers bought an expensive subscription to the firm's credit report services which led to salespeople fraudulently overstating their historical usage.
  - More recent examples of accounting manipulation include: Enron, Tyco and WorldCom, to name just three of the largest and most egregious cases.
- Suppose that the principal's benefit function b is given by $b = f_1 a_1 + f_2 a_2 + \epsilon$.
- But this cannot be contracted on.
- What can be contracted on is a performance measure $p = g_1 a_1 + g_2 a_2 + \phi$.
- Assume linear contracts of the form $w = t + vp$.
- Risk neutral P and A
- For simplicity assume that $E(\epsilon) = E(\phi) = 0$ and that $\psi(a_1, a_2) = \frac{1}{2}\left(a_1^2 + a_2^2\right)$.
- Agent therefore solves the following problem

$$\max_{a_1, a_2} \left\{ t + v\left(g_1 a_1 + g_2 a_2\right) - \frac{1}{2}\left(a_1^2 + a_2^2\right) \right\}.$$

- The agent's optimal actions are $a_1^*(v) = g_1 v$, and $a_2^*(v) = g_2 v$.
- What is the optimal level of v–i.e. what is the optimal intensity of incentives? Given the we have just worked out the optimal actions for the agent we can deduce that the principal's expected payoff is

$$E[b - w] = f_1 a_1^*(v) + f_2 a_2^*(v) - t - v g_1 a_1^*(v) - v g_2 a_2^*(v).$$

- A's expected payoff is

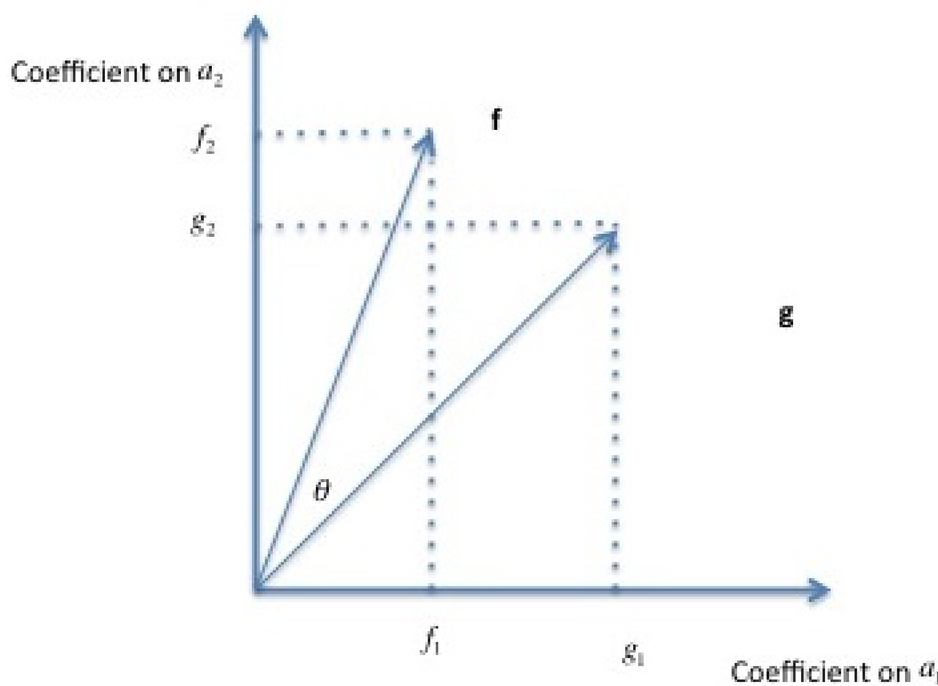
$$\begin{aligned} E[U] &= E[w] - \psi(a_1, a_2) \\ &= t + v\left(g_1 a_1^*(v) + g_2 a_2^*(v)\right) - \frac{1}{2}\left(a_1^*(v)^2 + a_2^*(v)^2\right). \end{aligned}$$

- Total surplus (the sum of the principal's and agent's payoff) is then

$$E[y] - \psi(a_1, a_2) = f_1 a_1^*(v) + f_2 a_2^*(v) - \frac{1}{2}\left(a_1^*(v)^2 + a_2^*(v)^2\right).$$

- It is easy to show that the optimal choice of v is then

$$v^* = \frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2}. \tag{30}$$



- First notice the importance of *scale*
  - Suppose both g_1 and g_2 were much larger than f_1 and f_2.
  - The agent could cause measurable performance, p, to go up by putting in lots of effort (high a_1 and a_2).
  - But this would lead to a much lower expected benefit for the principal.
  - So the efficient contract makes v–the incentive to make p high–very low.
- The second thing to note is *alignment*.
  - Suppose θ is small.

- Then a high value of v (powerful incentives) has a great effect in increasing b, the benefit to the principal.
- Now suppose that θ is large, so that $\mathbf{f}$ and $\mathbf{g}$ are poorly aligned.
- In this case incentives for increasing the performance measure (high v) are not very helpful for increasing b.
- Indeed, in the extreme case where the f and g vectors are orthogonal, incentives are useless in increasing b.

- With a little middle-school math we can derive

$$v^* = \frac{\sqrt{f_1^2 + f_2^2}}{\sqrt{g_1^2 + g_2^2}} \cos(\theta), \tag{31}$$

where θ is the angle between the f and g vectors.

- What makes for a good performance measure?
- It is tempting to say that if p is highly correlated with b then it is a good performance measure.
- One way in which this can be the case is if the noise terms, ϵ and ϕ are highly correlated.
- Does this make p a good performance measure?
- No. Imagine that p was the stock price and b was fundamental value. It's plausible to think that shocks to both are highly correlated, but there could be a wide divergence between the two measures if markets are not efficient in the short-run (as much evidence suggests is often the case).
- Though highly correlated with true value, the stock price would not be a good performance measure.
- A stark way to see this is to suppose that $p = a_1 + \epsilon$ and $b = a_2 + \epsilon$.
- The shocks are perfectly correlated, but p is a terrible performance measure.
- Put a different way, the only way to judge what is a good performance measure is through the actions it induces. This is what I call the *Gibbons tautology*: a performance measure is valuable if and only if it induces valuable actions. In other words, alignment is more important than noise.

5.8.4 Mixing Relational and Explicit Contracts

- Based on Baker-Gibbons-Murphy (QJE, 1994).
- Setup basically as above, but let $H = 1$ and $L = 0$.
- y cannot be contracted on.
- A's action also affect a contractible performance measure p (that is also 0 or 1).
- Before taking a, A privately observes μ.

- $\Pr(p=1) = \mu a$, and p and y are conditionally independent, and also assume that $E[\mu] = 1$ (on avg. p is an unbiased measure of μ).
- The implicit bonus is b and the explicit bonus is β that is paid when $p = 1$.
- Timing as above (with the objective bonus being paid after output is realized).
- Let $c(a) = \gamma a^2$.
- The FB has $a^{FB} = 1/2\gamma$.
- If A believes P will honor the implicit contract A solves

$$\max_a \left\{ s + ab + \mu a\beta - \gamma a^2 \right\}.$$

- The FOC is

$$a^* = \frac{b + \mu\beta}{2\gamma}.$$

- So we know that $a^* < a^{FB}$ if $b + \mu\beta < 1$.
- A will work for the firm if IR(A) is satisfied (before seeing μ.)
- So P's expected per-period profit , given the relational contract embodied in b and the explicit contract in β is:

$$E_\mu \left(a^*(\mu, b, \beta) - [s + a^*(\mu, b, \beta)b + \mu a^*(\mu, b, \beta)\beta]\right).$$

- P lowers s to make IR(A) binding (due to separability), so we have expected profit function for P:

$$V(b, \beta) = E_\mu \left(a^*(\mu, b, \beta) - \gamma a^*(\mu, b, \beta)^2 - \overline{w}\right)$$

- Now combine explicit and implicit
- At the end of each period P and A observe the objective performance measure p and also y.
- If $p = 1$ P pays bonus β according to the explicit contract and if $y = 1$ P chooses whether to pay A the bonus b specified in the RC.
- Before, P's expected profit if it reneged was zero in all future periods, because the trigger strategy implied that A wouldn't work for P.
- This is different when both explicit and implicit measures are available.
- One thing is that the expected per-period profit from honoring the RC is not $V(b)$ but $V(b, \beta)$–see above for the equations.
- Second thing, if P reneges, A would not accept any future RCs, but might want to take an explicit contract (if it was sufficiently remunerative).

- Without RCs, the per-period expected profit from (optimal) explicit contracting is $V(\beta^*)$ can be positive or negative.
- Depends on A's outside option $\overline{w}$ and how distorted the performance measure is $(var(\mu))$.
- If $V(\beta^*) > 0$–so explicit contracting can satisfy IR(A) and generate positive profits for P, the outside option for P if she reneges on the RC is $V(\beta^*) > 0$.
- But if $V(\beta^*) < 0$, the P's outside option is to shutdown and get zero.
- These outside options are important for understanding the optimal mix of implicit and explicit contracts.
- The reneging constraint, if $V(\beta^*) > 0$ is:

$$\frac{V(b,\beta) - V(\beta^*)}{r} \geq b \rightarrow V(b,\beta) - V(\beta^*) \geq rb.$$

- In the case where $V(\beta^*) > 0$ it can be shown (by solving for the optimal contract) that RCs cannot be (optimally) used when the discount rate is sufficiently high or the amount of distortion between the performance measure and output is sufficiently low.
  - If the performance measure is v. good then P's outside option is also v. good–so the temptation to renege is high.
  - A remarkable fact is that under parameter values that mean that an RC alone could get the FB, the very possibility of an imperfect explicit contract could render the RC infeasible.
- Also, we can get FB if the discount rate is sufficiently low.
- But even for a low discount rate FB cannot be achieved if the performance measure is nearly perfect, so the outside option in close to FB.
- Now suppose that $V(\beta^*) < 0$ so that P's outside option after reneging on an RC is to get zero profit by shutting gown the firm.
- This happens (naturally) when distortions caused by a "bad" performance measure are high.
- But could think of other ways this could happen (outside the model)
  - There are multiple implicit contracts, and destroying the reputation for one destroys all RCs
  - There are multiple recipients of RCs and ratting on one destroys the other RC
  - ...
  - Highlights the importance of beliefs about what the implicit contracting variables actually are
- Now the reneging constraint is not $V(b,\beta) - V(\beta^*) \geq rb$, but rather $V(b,\beta^*) \geq rb$.
- One can again solve for the optimal contract.

- A striking result is that as the performance measure becomes less distortionary, the implicit bonus *increases*, but the effect on the explicit bonus is ambiguous.
  - Suppose P and A have an optimal RC, but with no explicit contract.
  - Suppose also that the discount rate is high enough that the reneging constraint binds.
  - Now drop in a contractible performance measure, but one that is bad enough that it couldn't support a profitable explicit contract without an RC in the background.
  - It can still be useful.
  - A low-powered explicit contract can improve the future value of the relationship, so the reneging constraint no longer binds.
  - So the implicit bonus in the RC can be made larger.
- Here, implicit and explicit incentive are complements.
- Analogous result holds for *better* performance measures, not just the introduction of a new one.

5.8.5 Levin's Model

- Consider a sequence of spot contracts between a principal (P) and agent (A)
- Assume both are risk-neutral
- Assume both have common discount factor $\delta < 1$
- Let per period reservation utilities be $\bar{V}$ and $\bar{U}$ for P and A respectively and let $\bar{s} = \bar{V} + \bar{U}$
- A chooses action $a \in A$
- Output levels $q_1 < ... < q_n$
- Probability of these is $\pi_i(a)$ (just like in Grossman-Hart, where π is a mapping from A to the probability simplex)
- Denote action in period t as a_t
- Assume $\pi_i(a) > 0$ for all i and that MLRP holds
- Payment from P to A in period t is $I_t = w_t + b_t$ (interpreted as wage plus bonus)
- P's per period payoff is $q_i^t - I_t$
- A's per period payoff is $I_t - \psi(a_t, \theta_t)$, where θ_t is a cost parameter which is private information
- Let $\theta_t \in \{\theta_L, \theta_H\}$ with $\theta_L \leq \theta_H$
- Assume that these are iid over time and let $\beta = \Pr(\theta_t = \theta_H)$

- Assume ψ is convex, increasing and that $\psi(0,\theta)=0$, that $\psi_\theta(\cdot)>0$ and $\psi_{a\theta}(\cdot)>0$, where subscripts denote partial derivatives
- First best in a given period solves

$$\max_{a\in A}\left\{\sum_{i=1}^{n}\pi_i(a)\,q_i-\psi(a,\theta)\right\}$$

- Let

$$a^{FB}(\theta)=\arg\max_{a\in A}\left\{\sum_{i=1}^{n}\pi_i(a)\,q_i-\psi(a,\theta)\right\}$$

  and assume uniqueness
- Also assume

$$\sum_{i=1}^{n}\pi_i\left(a^{FB}\right)q_i-\psi\left(a^{FB},\theta\right)>\bar{s}$$

- Otherwise it would not be an interesting problem.
- Consider the game where each period the players choose whether or not to participate, A chooses an action and P chooses an output contingent bonus payment $b_t\left(q_i^t\right)$

Definition 19. *A Relational Contract is a perfect Bayesian equilibrium of the above game.*

- Let σ^A and σ^P be the strategy A and P respectively
- These are a function of observed history of play and output realizations
- Not contingent on A's action because it is not observable to P, and is sunk for A, and past actions don't affect P's continuation play.
- Assume that output realizations are observable but *not* verifiable
- Assume that past payments are observable *and* verifiable
- Let ζ^w be flow payoffs from verifiable components and ζ^b be from non-verifiable components
- ζ^b is the self-enforced part and it specifies a bonus payment $b_t(h_t)$, where h_t is the history of play and output realizations up to t

Definition 20. *We say that a relational contract is Stationary if in every period* $a_t=a(\theta_t), b_t=b\left(q_i^t\right)$ *and* $w_t=w$ *on the equilibrium path.*

- Levin (2003) proves that one can restrict attention to stationary contracts wlog
  - Basic argument is that for any set of non-stationary transfers and actions one can find a stationary contract with the same payoffs
  - Can't get joint punishment with a stationary contract–but it turns out that when P's behavior is observable, optimal contracts don't involve joint punishment in equilibrium

- Fix a relational contract $\left(\sigma^A, \sigma^P, \zeta^w, \zeta^b\right)$ and let $\hat{u}$ be A's payoff under this contract and $\hat{s} - \hat{u}$ be P's payoff
- Similarly, let $\hat{w}$ be the wage (which is court enforcable), $\hat{b}(q_i)$ be the bonus payment under this contract, and $\hat{a}(\theta)$ be A's action
- Joint value is then given by the program

$$\hat{s} = \max_{a(\theta)} \{(1-\delta) E_{\theta,q} [q - \psi(a(\theta), \theta) | a(\theta)] + \delta E_{\theta,q} [\hat{s}|\hat{a}(\theta)]\}$$

subject to

$$a(\theta) \in \arg\max_{a \in A} \left\{ E_q \left[\hat{w} + \hat{b}(q_i) + \frac{\delta}{1-\delta} \hat{u} | a \right] - \psi(a, \theta) \right\} \quad \text{(IC)}$$

$$\hat{b}(q_i) + \frac{\delta}{1-\delta} \hat{u} \geq \frac{\delta}{1-\delta} \bar{U} \quad \text{(IR-A)}$$

$$-\hat{b}(q_i) + \frac{\delta}{1-\delta} (\hat{s} - \hat{u}) \geq \frac{\delta}{1-\delta} \bar{V} \quad \text{(IR-P)}$$

- We are assuming that when A leaves the relationship she leaves forever (this is the strongest threat she has an gives rise to the largest set of relational contracts)–so no loss of generality.
- IR-P says P is willing to make the promised bonus payments (the RHS is what she gets if she doesn't pay and hence A quits forever; LHS is future payoff minus the bonus payment.)
- The contract which solves the program constitutes a PBE
  - If P doesn't participate at some point then P's best response is to not participate as well–and vice versa
- What about renegotiation?
  - Stationary contracts can be made renegotiation proof
    - If renegotiation precluded enforcement of inefficient punishments such that $s_\tau < \hat{s}$ after a deviation in period $t < \tau$, and RC can still be enforced with jointly efficient punishments by changing the split of the surplus.
    - P promises $\hat{w}$ and $\hat{b}(q_i)$ unless someone deviates.
    - If A deviates we move to a new contract with payments $\hat{\hat{w}} + \hat{\hat{b}}(q_i)$ such that $\hat{\hat{u}} = \overline{U}$ and $\hat{\hat{V}} = \hat{s} - \hat{\hat{u}}$ following the deviation.
    - If P deviates the contract has payments such that P gets $\overline{V}$.
    - If more deviations occur then do similarly.
    - Has to be renegotiation-proof because it's on the (constrained) Pareto frontier.
- What about existence?
  - It can be shown that a solution exists

  - When θ and q come from finite sets one can use the standard arguments from static Moral Hazard (e.g. Grossman-Hart 1983) and Adverse Selection (e.g. Mirrlees)

- Bonus payments can be positive or negative depending on how the surplus needs to be shared
  - If P gets "a lot" of the surplus then bonuses are positive–looks like incentive pay
  - Need to give big bonuses to satisfy IR-A when $\hat{u}$ is close to $\bar{u}$
  - If A gets "a lot" of the surplus then bonuses are negative–looks like efficiency wages
- Let $\bar{b}$ and $\underline{b}$ be the highest and lowest bonuses
- Then IR-A and IR-P combine to give the "self-enforcement constraint"

$$\left(\bar{b} - \underline{b}\right) \leq \frac{\delta}{1-\delta}\left(\hat{s} - \bar{s}\right)$$

- Can now compare relational contracts to contracts contractible output in the case of moral hazard
- Moral hazard (with no adverse selection) has $\theta_L = \theta_H = \theta$ which is common knowledge
- Risk-neutral P and A so optimal contract involves making A the residual claimant
- The payment scheme is

$$I = q_i + \bar{u} - \max_{a \in A}\left\{E_q\left[q|a\right] - \psi\left(a, \theta\right)\right\}$$

- This will violate the self-enforcement constraint if

$$\left(q_n - q_1\right) > \frac{\delta}{1-\delta}\left(E_q\left[q|a^{FB}\right] - \psi\left(a^{FB}, \theta\right) - \bar{s}\right)$$

- It can be shown that when this is violated the optimal relational contract is of the following form

$$\begin{aligned} b\left(q_i\right) &= \bar{b} \text{ for } q_i \geq q_k \\ b\left(q_i\right) &= \underline{b} \text{ for } q_i < q_k \end{aligned}$$

  where q_k is some interior cutoff value
- MLRP important here
- Connection to one period moral hazard with risk-neutrality and limited liability constraint (Innes, 1990)–SEC constraint acts like LL constraint.
- Can also apply the model to the case of pure adverse selection
  - That corresponds to a being observable to P and A, but θ being A's private information

- Note that Spence-Mirrlees condition is satisfied because we assumed that $\psi_{a\theta} > 0$. That condition is

$$\frac{\partial}{\partial\theta}\left(\frac{\partial[I-\psi(a,\theta)]}{\partial a}\right) < 0.$$

- Can be shown that the no distortion for the highest type no longer applies in the relational model–second-best enforceable contracts can violate SEC
  - The bonus payments in the court enforceable model can violate the self-enforcement constraint
  - So all types under-provide "effort"
  - Also get bunching with many types
- A general point–the self-enforcement constraint lowers the power of the incentives that can be provided (in either setting)
- Can also extend the model (as Levin does) to *subjective* performance measures
  - Stationary contracts now have problems
  - But the optimal contract is still quite simple
  - P pays A a base salary each period, and then a bonus if P (subjectively) judges performance to be above a threshold
  - But if below threshold then the relationship terminates
  - Inefficiency can come from the different beliefs about performance
  - So a mediator can be thought of as making the information more objective and therefore reducing the welfare loss
  - Can do better by making evaluation less frequent–can allow P to make more accurate assessments

5.8.6 Building Routines

- Based on Sylvain Chassang (AER, 2010).
- Basic question: suppose the details of cooperation are not common knowledge, how do players learn those details.
- Specifically: repeated game setting where one player has incomplete information about the timing and ability of the other player to affect her outcome.
- Imperfect monitoring, that at the start requires punishments that are inefficient.
- But as the common history grows there is a reduction in inefficiency–interpret as "better routines".
- Think back to Bull (1987)–we made relational contracts sound rather straightforward (e.g. subjective bonuses)
- How do parties specify what the contingencies *are* in a relational contract?!

- Link to private monitoring literature (Green-Porter 1984, Abreu-Pearce-Stachetti 1990, Fudenberg-Levine-Maskin 1994: all ECMA). See Kandori (JET, 2002) for an excellent overview: `http://personal.lse.ac.uk/zapal/EC501_2008_2009/Skrzypacz_background1.pdf`
- 2 players in a repeated game.
- Each period:
  - Player 1 decides whether to stay in the relationship and interact with player 2, or skip one period.
  - If player 1 stays, then player 2 gets to take an action.
  - The available actions are randomly drawn from some subset of the total action set.
  - Two types of actions: *productive actions* that give player 1 a benefit with positive probability, but also fail with some probability, and *unproductive actions* that cost player 2 zero, but do nothing for player 1.
  - At the end of each period the: set of available actions, the action P2 took, and the benefit P1 got are commonly observed.
  - Player 2 knows which actions are productive, but player 1 does not.
- The informational asymmetry means that P1 lacks the ability to interpret what the actions that P2 has available and took mean for payoffs.
- Wedge b/w the availability of information and the ability to interpret it.
- Basic idea: P1 doesn't know which actions are productive and thus when she should expect P2 to take a productive action (i.e. cooperate).
- So if P1 sees an action that has no benefit for her she doesn't know whether if P2 took a productive action and it failed, or if P2 chose and unproductive action (at zero cost to P2).
- So P1 might use inefficient exit in equilm (c.f. Green-Porter) to support cooperation.
- But once an action (say a_0) yields a benefit to P1, P1 knows that it's a productive action–so the monitoring problem goes away.
- Tradeoff between new productive actions (virtue: more beneficial, vice: monitoring problem) and existing actions that are known to be productive (virtue: no monitoring problem, vice: not as productive).
- Focus on Pareto efficient, pure strategy PBEa.
- Main results
  - During the "specification phase" the relationship is sensitive to shocks.
  - Indeed, during this phase, a P2 action that fails is followed by punishment from P1 on the equilm path.
  - But once the learning phase is done (we're in a routine) this is no longer the case.

- Can get path-dependence: costs of information revelation can be bigger than the efficiency gains from using a better routine.
- Ex ante identical partnerships can have different long-run outcomes because of initial shock realizations.

5.9 Markets and Contracts

5.9.1 Overview

- A lot of what we have done thus far considers bi-lateral (or sometimes multilateral) relationships
- But in some/many contexts, contracts between agents exist in market settings
- This has been recognized for a long time–Rothschild and Stiglitz (1976) analyze screening in such a context
- But there are a number of other issues of interest
- We will only touch on a few of them here
  - Contracts as a Barrier to Entry
  - Product Market Competition
  - Equilibrium Incentive Contracts

5.9.2 Contracts as a Barrier to Entry

- There is a long traditional in legal scholarship/law and economics which argues that contracts can be anti-competitive in effect
- Sellers may be able to "lock up" buyers with long-term contracts which prevent or at least deter entry to some degree
- Key reference is Aghion and Bolton (1987)
- Contracts that specify penalties for early termination can be used to extract rents from future entrants who may be lower cost than the current provider
- Suppose there are two time periods $t = 1$ and $t = 2$
- At $t = 1$ there is an incumbent who can sell a product at cost $c_I \leq 1/2$ and a buyer has reservation value $v = 1$ for this widget
- At $t = 2$ a potential entrant has cost c_E which is uniformly distributed on $[0, 1]$
- Obviously $p_1 = 1$ in period 1
- Assume that if entry occurs there is Bertrand competition at $t = 2$
- So entry occurs if $c_E \leq c_I$
- If there is no contract / a spot contract then if entry occurs $p_2 = \max\{c_E, c_I\} = c_I$ and if no entry then $p_2 = 1$

- So under the spot contract the expected payoff of the buyer is

$$\begin{aligned} V_B &= (1 - \Pr(entry))\,0 + \Pr(entry)(1 - c_I) \\ &= c_I\,(1 - c_I) \end{aligned}$$

- And the incumbent firm's payoff is

$$\begin{aligned} V_I &= p_1 - 1 + (1 - \Pr(entry))\,(1 - c_I) + \Pr(entry)\,(1 - c_I) \\ &= 1 - c_I + (1 - c_I)^2 \end{aligned}$$

- Now consider the case where the incumbent and the buyer sign a contract at $t = 1$ which specifies as price for each period and a penalty d for breach / termination
  - The contract is a triple (p_1, p_2, d)
- So the buyer will only breach the contract if the entrants price p_E is such that

$$1 - p_E \geq 1 - p_2 + d$$

  i.e. surplus under the new contract compensates for the surplus under the old including damages

- The probability of entry given this contract is

$$\Pr(c_E < p_2 - d) = p_2 - d$$

- The buyer's expected payoff under the contract is

$$\begin{aligned} V_B^L &= (1 - p_1) + (1 - p_E) + d \\ &= (1 - p_1) + (1 - (p_2 - d)) + d \\ &= (1 - p_1) + (1 - p_2) \end{aligned}$$

- The incumbent's expected payoff is

$$\begin{aligned} V_I^C &= p_1 - c_I + (1 - \Pr(entry))\,(p_2 - c_I) + \Pr(entry)d \\ &= p_1 - c_I + (1 - p_2 + d)\,(p_2 - c_I) + (p_2 - d)\,d \end{aligned}$$

- The buyer will only accept the contract if

$$(1 - p_1) + (1 - p_2) \geq c_I\,(1 - c_I)$$

- So the incumbent solves

$$\max_{p_1, p_2, d} \{p_1 - c_I + (1 - p_2 + d)\,(p_2 - c_I) + (p_2 - d)\,d\}$$

$$\text{subject to}$$

$$(1 - p_1) + (1 - p_2) \geq c_I\,(1 - c_I)$$

i.e. maximize the payoff under the contract subject to the buyer being willing to accept

- The incumbent can always set $p_1 = 1$, so the problem is

$$\max_{p_2,d} \{1 - c_I + (1 - p_2 + d)(p_2 - c_I) + (p_2 - d)\, d\}$$

subject to

$$(1 - p_2) \geq c_I (1 - c_I)$$

- Noting that the constraint binds we have $1 - c_I (1 - c_I) = p_2$
- So the program is

$$\max_{d} \{1 - c_I + (1 - (1 - c_I (1 - c_I)) + d)((1 - c_I (1 - c_I)) - c_I) + ((1 - c_I (1 - c_I)) - d)\, d\}$$

- The solution is

$$d^* = \frac{1 + (1 - c_I)(1 - 2c_I)}{2} > 0$$

- So the probability of entry is

$$p_2 - d^* = \frac{c_I}{2}$$

- The incumbent always wants to sign the contract
- This contract is competition reducing since the probability of entry is $\frac{c_I}{2}$ instead of c_I
- Markets with contracts may not be as efficient as spot contract markets!
- Robust to certain extensions
  - Renegotiation
  - Multiple buyers

5.9.3 Multiple Principal-Agent Dyads

- We have focused on a number of models of a single principal and a single agent.
- Situations with several principals but one agent are quite prevalent (e.g. wholesale trade, politics).
- This is referred to as *common agency*–see Bernheim-Whinston (ECMA, 1986).
- Won't focus on that here.
- Will focus on there are multiple P-A dyads.
- Also important for understanding oligopoly–since much (most?) competition takes place between "managerial firms".
- Do contracts between principals and agents affect the nature of competition, and are there strategic effects?

- Key early paper is Fershtman-Judd (AER, 1987).
- Suppose that firm i consists of a shareholder (principal) and a manager (agent).
- Both are risk-neutral.
- Let output be q_i and profits be Π_i.
- Time 1: each dyad contracts simultaneously on a linear incentive scheme of the form $w_i = \alpha_i q_i + \beta_i \Pi_i$.
- Time 2: each dyad's contract becomes public knowledge and then the firms compete in the product market: (differentiated products) Bertand or Cournot.
- Under Cournot competition there is strategic substitutability (reaction functions slope down).
- Would like to commit to being tough because this induces competitors to act softer.
- One way to do this is to provide incentives for *output*–i.e. $\alpha_i > 0$.
- Under Bertrand competition there is strategic complementarity, so it is optimal to set $\alpha_i < 0$.
- $\alpha_i = 0$ when there is a product market in which there is no substitutability or complementarity.
- Issues:
  - Predictions depend crucially on the nature of product market competition–easy to write down in models, but harder to determine empirically. Is the world Cournot?
  - Renegotiation: each dyad would like to fool other dyads into believing that incentives for output are there, but they are not profit maximizing conditional on that belief.
  - Secret renegotiation with symmetric information implies $\alpha_i = 0$ for all i.
- Can introduce asymmetric information within each dyad to prevent perfect renegotiation (link to renegotiation in classic P-A model: Fudenberg-Tirole (ECMA, 1990), Hermalin-Katz (ECMA, 1991).
- Caillaud-Julien-Picard (ECMA, 1995) model secret renegotiation in this setting (and agents have different cost types, not known at the time of contracting).
- Two firms.
- Time 1: initial contract about R. Time 2: secret renegotiation and contract offered on output of the two firms $t(a, b)$. Time 3: agent learns type and output occurs (or agent can leave the market).
- Roughly, they find that under Cournot competition signing an initial contract with some unconditional payment R, is optimal.

- Equilm initial contract gives agent rents that limit the need to reduce output below the full-information Cournot level.
- This shifts the firm's reaction function up and hence lowers the other firm's output.
- Under differentiated products Bertrand this is not true.
- Strategic effect of initial contract R is bad–makes the other firm tougher.
- So no initial contract.
- Basic point: under Cournot the initial contract relaxes the IR constraint in an adverse selection setting.

5.9.4 Product Market Competition and the Principal-Agent Problem

- Classic question: does product market competition increase internal efficiency of the firm?
- Leibenstein (1967): internal firm inefficiency–"X-Inefficiency"–may be very large
- Does competition help?
- Hicks (1935): "The best of all monopoly profits is a quiet life"
- First formal model is Hart (1983)–satisfising behavior
- Scharfstein (1987) with Hart's model but different utility function obtains opposite conclusion
- Martin (1993)–Cournot competition means less effort
- Many others–see Holden (2005) for references
- Will focus on three models due to Schmidt (1997)
- Look at these through the lens of Holden (2005) framework
- There are two players, a risk-neutral principal and a risk-averse agent
- Let $\phi \in \mathbb{R}$ be a measure of product market competition which affects the profits which accrue to the principal.
- A higher value of ϕ means that, all else equal, profits are lower.
- Suppose that there are a finite number of possible gross profit levels for the firm. Denote these $q_1(\phi) < ... < q_n(\phi)$.
- These are profits before any payments to the agent.

Definition 21. *A set X is a* ***product set*** *if $\exists$ sets $X_1, ..., X_n$ such that $X = X_1 \times ... \times X_n$. X is a* ***product set in*** *$\mathbb{R}^n$ if $X_i \subseteq \mathbb{R}, i = 1, ..., n$.*

- The set of actions available to the agent, A, is assumed to be a product set in $\mathbb{R}^n$ which is closed, bounded and non-empty.

- Let S be the standard probability simplex, i.e. $S = \{y \in \mathbb{R}^n | y \geq 0, \sum_{i=1}^n y_i = 1\}$ and assume that there is a twice continuously differentiable function $\pi : A \to S$. The probabilities of outcomes $q_1(\phi), ..., q_n(\phi)$ are therefore $\pi_1(a), ..., \pi_n(a)$.
- Let the agent's von Neumann-Morgenstern utility function be of the following form:

$$U(a, I) = G(a) + K(a)V(I)$$

where I is a payment from the principal to the agent, and $a \in A$ is the action taken by the agent.

Definition 22. *An **incentive scheme** is an n-dimensional vector* $\mathbf{I} = (I_1, ..., I_n) \in \mathcal{I}^n$.

- Given an incentive scheme the agent chooses $a \in A$ to maximize her expected utility $\sum_{i=1}^n \pi_i(a) U(a, I_i)$.
- Key condition for increase in product market competition to decrease agency costs is

$$\sum_{i=1}^{n} q_i'(\phi)\pi_i'(a) \geq 0, \forall a, \phi. \tag{32}$$

- When MLRP holds this become

$$\sum_{i=j+1}^{n} \pi_i'(a) q_i'(\phi) \geq \sum_{i=1}^{j} |\pi_i'(a)| \, q_i'(\phi). \tag{33}$$

Schmidt's Basic Model

- The firm goes bankrupt if realized profits are below a certain level
- Reduced form measure of product market competition, ϕ
- An increase in ϕ corresponds to a more competitive product market
- Effort by the agent affects costs
- Two possible states: high cost and low cost–states L and H
- (33) becomes:

$$\pi_L'(a) \left[q_L'(\phi) - q_H'(\phi)\right] > 0 \tag{34}$$

- By FOSD $\pi_L'(a) > 0$ (a harder action makes the low cost state more likely)
- Schmidt's result requires $q_H'(\phi) < q_L'(\phi)$
- True because loss on the agent of $\overline{L}$ if the firms goes bankrupt
  - Occurs with positive probability in the high cost state and with zero probability in the low cost state
  - He assumes that the probability of this occurring is $l(\phi)$ with $l'(\phi) > 0$

- This loss of $\overline{L}$ is equivalent to profits being lower since it affects the agent's utility and hence the payment that the Principal must make if the participation constraint binds
- In effect, then $q_H(\phi) \equiv \overline{q}_H(\phi) - l(\phi)\overline{L}$
- Schmidt's main result states that the increase in agent effort is unambiguous if the PC binds
- In such circumstances $q_L'(\phi) > q_H'(\phi)$, since the expected loss of $\mathbb{E}[L]$ occurs only in state H
- If the PC is slack at the optimum then the effect of competition is ambiguous because the loss of L is only equivalent to profits being lower if L is sufficiently large
- Thus, for $\overline{L}$ sufficiently small we have $q_L'(\phi) = q_H'(\phi)$ and hence the condition is not satisfied.

Schmidt's Price-Cap Model

- Now consider price-cap regulation of a monopolis
- Firm can have constant marginal cost of either c^L or $c^H > c^L$
- Regulator does not observe costs, but sets a price cap of $1/\phi$
- Larger value of ϕ interpreted as a more competitive product market.
- Denoting demand at the cap (which is assumed to be binding regardless of the cost realization) as $D(1/\phi)$, profits are:

$$q(c^j, \phi) = D\left(\frac{1}{\phi}\right)\left(\frac{1}{\phi} - c^j\right)$$

- Differentiating with respect to ϕ yields:

$$\frac{\partial q(c^j, \phi)}{\partial \phi} = -\frac{1}{\phi^2}\left[D\left(\frac{1}{\phi}\right) + D'\left(\frac{1}{\phi}\right)\left(\frac{1}{\phi} - c^j\right)\right]$$

- General condition for a harder action in this two outcome model is simply:

$$\pi_L'(a)\left[q_L'(\phi) - q_H'(\phi)\right] \geq 0$$

- Since $\pi_L'(a)$ is positive, we require $q_L'(\phi) - q_H'(\phi) \geq 0$ – i.e. $q_L'(\phi) \geq q_H'(\phi)$. This requires:

$$-\frac{1}{\phi^2}\left[D\left(\frac{1}{\phi}\right) + D'\left(\frac{1}{\phi}\right)\left(\frac{1}{\phi} - c^L\right)\right] \geq$$
$$-\frac{1}{\phi^2}\left[D\left(\frac{1}{\phi}\right) + D'\left(\frac{1}{\phi}\right)\left(\frac{1}{\phi} - c^H\right)\right]$$

- which reduces to requiring:

$$\frac{(c^L - c^H)D'\left(\frac{1}{\phi}\right)}{\phi^2} \geq 0$$

  Obviously $D'\left(\frac{1}{\phi}\right) < 0$, and, by construction, $c^H > c^L$.

- A tighter price cap leads to a harder action by the agent.

5.9.5 Equilibrium Effort Effects

Definition 23. *A Noncooperative game is a triple* $(N, \mathbf{S}, \{\mathbf{f}_i \colon \mathbf{i} \in \mathbf{N}\})$, *where* N *is a nonempty, finite set of players,* S *is a set of feasible joint strategies,* $f_i(x)$ *is the payoff function for player* i, *which is real-valued on* S, *a strategy for each player* i *is an* m_i *vector* x_i, *and a joint strategy is an* $\{x_i : i \in N\}$.

Definition 24. *A noncooperative game* $(N, \mathbf{S}, \{\mathbf{f}_i \colon \mathbf{i} \in \mathbf{N}\})$, *is a Supermodular Game if the set* S *of feasible joint strategies is a sublattice of* $\mathbb{R}^m$, *the payoff function* $f_i(\mathbf{y}_i, \mathbf{x}_{-i})$ *is supermodular in* y_i *on* S_i *for each* x_{-i} *in* S_{-i} *and each player* i, *and* $f_i(\mathbf{y}_i, \mathbf{x}_{-i})$ *has increasing differences in* $(\mathbf{y}_i, \mathbf{x}_{-i})$ *on* $S_i \times S_{-i}$ *for each* i.

Theorem 13 (Topkis 4.2.3)**.** *Suppose that* $(N, \mathbf{S}, \{\mathbf{f}_i \colon \mathbf{i} \in \mathbf{N}\})$ *is a supermodular game, the set* S *of feasible joint strategies in nonempty and compact, and the payoff function* $f_i(\mathbf{y}_i, \mathbf{x}_{-i})$ *is upper semicontinuous in* y_i *on* $S_i(\mathbf{x}_{-i})$ *for each player* i *and each* x_{-i} *in* S_{-i}. *For each* x *in* S *and each subset* N' *of* N, *let* $x_{N'} = \{x_i : i \in N'\}$. *Let* x' *be the least element of* S. *For each subset* N' *of* N, *let* $S^{N'}$ *be the section of* S *at* $x'_{N \setminus N'}$. *For each subset* N' *of* N, *each player* i *in* N', *and each* $x_{N'}$ *in* $S^{N'}$, *let* $f_i^{N'}(x_{N'}) = f_i(x_{N'}, x'_{N \setminus N'})$. *Consider the collection of supermodular games* $(N', S^{N'}, \{f_i^{N'} : i \in N'\})$ *parameterized by the nonempty subsets* N' *of* N. *Then there exists a greatest equilibrium point and a least equilibrium point for each game* N', *and for each player* i *the strategy of player* i *in the greatest (least) equilibrium point for game* N' *is increasing in* N' *where* i *is included in* N'.

- Topkis Theorem 4.2.3 provides conditions under which the strategy of *each* player in the greatest equilibrium point, and the least equilibrium point, is increasing in a parameter, t
- These two Theorems apply to a finite number of players
- But analogous results have been proved for infinitely many players–and also for quasi-supermodular games (see Milgrom and Shannon, 1996)
- Want to know conditions under which the principal of *every* firm in the market induces a harder action from her agent in the greatest and least equilibrium of the game
- Interpret a player as being a principal, and a strategy for her as being a feasible section-best action (correspondence), $a^{**} = \sup_{a \in A} \{B(a) - C(a)\}$, and a product market strategy $\mathbf{z}_i \in Z_i$, where Z_i is the set of product market strategies for player i

- If this game is a supermodular game then Topkis's theorems imply that the actions implemented by all principals are increasing in the relevant measure of product market competition
- First we need the set of feasible joint strategies be compact
- If the sets of product market strategies Z_i are nonempty and compact for all i then it follows trivially from Tychonoff's Theorem that the set S of feasible joint strategies in the Product Market with Agency Game is compact.
- e.g. if a product market strategy is a price, quantity or supply function then S will be compact.
- Second requirement: the payoff function is supermodular in $\mathbf{y}_i \in S_i$.
- The key part of this requirement is that the agent's action and the product market strategy be complements
- e.g. in a Cournot game where agent effort reduces cost this condition requires that lower costs make choosing higher quantities more desirable
- Whether or not this condition is met clearly depends on the nature of the product market and the effect of the agents' actions.
- The final important condition is that the payoff exhibit increasing differences in $(\mathbf{y}_i, \mathbf{x}_{-i})$ on $S_i \times S_{-i}$ for all i.
- Also depends on the particulars of the game.
- e.g. in Cournot, this requires that a higher effort-quantity pair from one firm makes a higher effort-quantity pair from another firm more desirable.