renegotiation can be made conditional on a. Thus, the main differences are whether renegotiation takes place between asymmetrically or (sufficiently) symmetrically informed parties.

3 Mechanism Design and Self-selection Contracts

3.1 Mechanism Design and the Revelation Principle

We consider a setting where the principal can offer a mechanism (e.g., contract, game, etc.) which her agents can play. The agent's are assumed to have private information about their preferences. Specifically, consider I agents indexed by $i \in \{1, \ldots, I\}$.

- Each agent i observes only its own preference parameter, $\theta_i \in \Theta_i$. Let $\theta \equiv (\theta_1, \ldots, \theta_I) \in \Theta \equiv \prod_{i=1}^{I} \Theta_i$.
- Let $y \in Y$ be an allocation. For example, we might have $y \equiv (x, t)$, with $x \equiv (x_1, \ldots, x_I)$ and $t \equiv (t_1, \ldots, t_I)$, and where x_i is agent i's consumption choice and t_i is the agent's payment to the principal. The choice of y is generally controlled by the principal, although she may commit to a particular set of rules.
- Utility for i is given by $U_i(y, \theta)$; note general interdependence of utilities on θ_{-i} and y_{-i}. The principal's utility is given by the function $V(y, \theta)$. In a slight abuse of notation, if y is a distribution of outcomes, then we'll let U_i and V represent the value of expected utility after integrating with respect to the distribution.
- Let $p(\theta_{-i}|\theta_i)$ be i's probability assessment over the possible types of other agents given his type is θ_i and let $p(\theta)$ be the common prior on possible types.

Suppose that the principal has all of the bargaining power and can commit to playing a particular game or mechanism involving her agent(s). Posed as a mechanism design question, the principal will want to choose the game (from the set of all possible games) which has the best equilibrium (to be defined) for the principal. But this set of all possible games is enormous and complex. The revelation principle, due to Green and Laffont [1977], Myerson [1979], Harris and Townsend [1981], Dasgupta, Hammond, and Maskin [1979], et al., allows us to simplify the problem dramatically.

Definition: A communication *mechanism* or game,

$$\Gamma^c \equiv \{\mathcal{M}, \Theta, p, U_i(y(m), \theta)_{i=1,\ldots,I}\},$$

is characterized by a message (i.e., strategy) space for each agent, $\mathcal{M}_i$, and an allocation y for each possible message profile, $m \equiv (m_1, \ldots, m_I) \in \mathcal{M} \equiv (\mathcal{M}_1, \ldots, \mathcal{M}_I)$; i.e., $y : \mathcal{M} \mapsto Y$. For generality, we will suppose that $\mathcal{M}_i$ includes all possible mixtures over messages; thus, m_i may be a probability distribution. When no confusion would result, we sometimes indicate a mechanism Γ by the pair $\{\mathcal{M}, y\}$.

The timing of the communication mechanism game is as follows:

- Stage 1. The principal offers a communication mechanism and a Nash equilibrium to play.
- Stage 2. The agents simultaneously decide whether or not to participate in the mechanism. (This stage may be superfluous in some contexts; moreover, we can always require the principal include the message of "I do not wish to play" and the null contract, making the acceptance stage unnecessary.)
- Stage 3. Agents play the communication mechanism.

The idea here is that the principal commits to a mechanism, $y(m)$, and the agents all choose their messages in light of this. To state the revelation principle, we must choose an equilibrium concept for Γ^c. We first consider Bayesian-Nash equilibria (other possibilities include dominant-strategy equilibria and correlated equilibria).

3.1.1 The Revelation Principle for Bayesian-Nash Equilibria

Let $m^*(\theta) \equiv (m_1^*(\theta_1), \ldots, m_I^*(\theta_I))$ be a Bayesian-Nash equilibrium (BNE) of the game in stage 3, and suppose without loss of generality that all agent's participated at stage 2. Then $y(m^*(\theta))$ denotes the equilibrium allocation.

Revelation Principle (BNE): Suppose that a mechanism, Γ^c has a BNE $m^*(\theta)$ defined over all θ which yields allocation $y(m^*(\theta))$. Then there exists a direct revelation mechanism, $\Gamma^d \equiv \{\mathcal{M}{\equiv}\Theta, \Theta, p, U_i(\tilde{y}(\theta), \theta)_{i=1,\ldots,I}\}$, with strategy spaces $\mathcal{M}_i \equiv \Theta_i, i = 1, \ldots, I$, and an outcome function $\tilde{y}(\theta) : \Theta \mapsto Y$ such that there exists a BNE in Γ^d with $\tilde{y}(\theta) = y(m^*(\theta))$ and equilibrium strategies $m_i(\theta_i) = \theta_i \; \forall \theta$.

Proof: Because $m^*(\theta)$ is a BNE of Γ^c, for any i with type θ_i,

$$m_i^*(\theta_i) \in \arg \max_{m_i \in \mathcal{M}_i} E_{\theta_{-i}}[U_i(y(m_i, m_{-i}^*(\theta_{-i})), \theta_i, \theta_{-i})|\theta_i].$$

This implies for any θ_i

$$E_{\theta_{-i}}[U_i(y(m_i^*(\theta_i), m_{-i}^*(\theta_{-i})), \theta_i, \theta_{-i})|\theta_i] \geq E_{\theta_{-i}}[U_i(y(m_i^*(\hat{\theta}_i), m_{-i}^*(\theta_{-i})), \theta_i, \theta_{-i})|\theta_i], \; \forall \hat{\theta}_i \in \Theta_i.$$

Let $y(\theta) \equiv y(m^*(\theta))$. Then we can rewrite the above equation as:

$$E_{\theta_{-i}}[U_i(y(\theta_i, \theta_{-i}), \theta_i, \theta_{-i})|\theta_i] \geq E_{\theta_{-i}}[U_i(y(\hat{\theta}_i, \theta_{-i}), \theta_i, \theta_{-i})|\theta_i], \; \forall \hat{\theta}_i \in \Theta_i.$$

But this implies $m_i(\theta_i) = \theta_i$ is an optimal strategy in Γ^d and $y(\theta)$ is an equilibrium allocation. Therefore, truthtelling is a BNE in the direct mechanism game. □

Remarks:

1. This is an extremely useful result. If a game exists in which a particular allocation y can be implemented by the principal, there is a direct revelation mechanism with truth-telling as an equilibrium that can also accomplish this. Hence, *without loss of generality, the principal can restrict attention to direct revelation mechanisms in which truth-telling is an equilibrium.*

2. The more general formulations of this principle such as by Myerson, allow agents to take actions as well. That is, y has some components which are under the agents' control and some components which are under the principal's control. A revelation principle still holds in which the principal implements y by choosing the components it controls and making suggestions to the agents as to which actions they should take that are under their control. Truthful revelation occurs in equilibrium and suggestions are followed. Myerson refers to this as "truthtelling" and "obedience," respectively.

3. This notion of truthful implementation in a BNE is a very weak concept. There may be many other equilibria to Γ^d which are not truthful and in which the agents do better. Thus, there may be strong reasons to believe that the agents will not follow the equilibrium the principal selects. This non-uniqueness problem has spawned a large number of papers which focus on conditions for a unique equilibrium allocation which are discussed in the survey articles by Moore [1992] (regarding symmetrically informed agents) and Palfrey [1992] (regarding asymmetrically informed agents).

4. We rarely see direct revelation mechanisms being used. Economically, the indirect mechanisms are more interesting to study once we find the direct mechanism. Possible advantages from carefully choosing an indirect mechanism are uniqueness, simplicity, and robustness against collusion, etc.

5. The key to the revelation principle is commitment. With commitment, the principal can replicate the outcome of any indirect mechanism by promising to play the strategy for each player that the player would have chosen in the indirect mechanism. Without commitment, we must be careful. Thus, when renegotiation is possible, the revelation principle fails to apply.

6. In some settings, agents contract with several principal's simultaneously (common agency), so there may be one agent working for two principals where each principal has control over a component of the allocation, y. Is there a related revelation principle such as "for any BNE in the common agency game with one agent and two principals, there exists a BNE to a pair of direct-revelation mechanisms (one offered by each principal) in which the agent reports truthfully to both principals"? The answer is no. The problem is that out-of-equilibrium messages which had no use in a one-principal setting, may enlarge the set of equilibria in the original game beyond those sustainable as equilibria to the revelation game in a multi-principal setting.

7. Note we could have just as easily used correlated equilibria or dominant-strategy equilibria as our equilibrium notion. There are similar revelation principles for these concepts.

3.1.2 The Revelation Principle for Dominant-Strategy Equilibria

If the principal wants to implement an allocation, y, in dominant strategies, then she has to design a mechanism such that this mechanism has a dominant-strategy equilibrium (DSE), $m^*(\theta)$, with outcome $y(m^*(\theta))$.

Revelation Principle (DSE): Suppose that Γ^c ha a dominant-strategy equilibrium, $m^*(\theta)$ with outcome $y(m^*(\theta))$. Then there exists a direct revelation mechanism, $\Gamma^d \equiv \{\Theta, y\}$, with strategy spaces, $\mathcal{M}_i \equiv \Theta_i,\ i = 1, \ldots, I$, and an outcome function $y(\theta) : \Theta \mapsto Y$ such that there exists a DSE in Γ^d in which truth-telling is a dominant strategy with DSE allocation $y(\theta) \equiv y(m^*(\theta))\ \forall \theta \in \Theta$.

Proof: Because $m^*(\theta)$ is a DSE of Γ^c, for any i and type $\theta_i \in \Theta_i$,

$$m_i^*(\theta_i) \in \arg\max_{m_i \in \mathcal{M}_i} U_i(y(m_i, m_{-i}), \theta_i, \theta_{-i}),\ \forall \theta_{-i} \in \Theta_{-i} \text{ and } \forall m_{-i} \in \mathcal{M}_{-i}.$$

This implies that $\forall\ (\hat{\theta}_i, \theta_{-i}) \in \Theta$

$$U_i(m_i^*(\theta_i), m_{-i}^*(\theta_{-i}), \theta_i, \theta_{-i}) \geq U_i(y(m_i^*(\hat{\theta}_i), m_{-i}^*(\theta_{-i})), \theta_i, \theta_{-i}).$$

Let $y(\theta) \equiv y(m^*(\theta))$. Then we can rewrite the above equation as

$$U_i(y(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}) \geq U_i(y(\hat{\theta}_i, \theta_{-i}), \theta_i, \theta_{-i}),\ \forall \hat{\theta}_i \in \Theta_i,\ \forall \theta_{-i} \in \Theta_{-i}.$$

But this implies truthtelling, $m_i(\theta_i) = \theta_i$, is a DSE of Γ^d with equilibrium allocation $y(\theta)$. $\square$

Remarks:

1. Certainly this is a more robust implementation concept. Dominant strategies are more likely to be played. Additionally, if for whatever reasons you believe that the agents have different priors, then the allocation is unchanged.

2. Generically, DSE are unique, although some economically likely environments are non-generic.

3. DSE is a (weakly) stronger concept than using BNE. But when only one agent is playing, the two concepts coincide.

4. We will generally focus in the BNE revelation principle, although we will discuss the use of dominant strategy mechanisms in a few simple settings. Furthermore, Mookerjee and Reichelstein [1992] have demonstrated that under a large class of contracting environments, the outcomes implemented with BNE mechanisms can be implemented with DSE mechanisms as well.

3.2 Static Principal-Agent Screening Contracts

With the revelation principle(s) developed, we proceed to characterize the set of truthful direct-revelation mechanisms and then to find the optimal mechanism from this set in a simple single-agent setting. We proceed first by exploring a simple two-type example of non-linear pricing, and then a detailed examination of the general case.

3.2.1 A Simple 2-type Model of Nonlinear Pricing

A risk-neutral firm produces a product of quality q at a cost per unit of $c(q)$. Its profit from a sale of one unit with quality q for a price of t is is $V = t - c(q)$. There are two types of consumers, $\underline{\theta}$ and $\overline{\theta}$, with $\overline{\theta} > \underline{\theta}$ and a proportion of p of type $\overline{\theta}$. For this example, we will assume that p is sufficiently small that the firm prefers to sell to both types rather than focus only on the the $\overline{\theta}$-type customer. Each consumer has a unit demand with utility from consuming a good of quality q for a price of t equal to $U = \theta q - t$, where $\theta \in \{\underline{\theta}, \overline{\theta}\}$.

By the revelation principle, the firm can restrict attention to contracts of the form $\{(\underline{q}, \underline{t}), (\overline{q}, \overline{t})\}$ such that type $\underline{t}$ consumers find it optimal to choose the first contract pair and $\overline{t}$ consumers choose the second pair. Thus, we can write the firm's optimization program as:

$$\max_{\{(\underline{q},\underline{t}),(\overline{q},\overline{t})\}} \; p[\overline{t} - c(\overline{q})] + (1-p)[\underline{t} - c(\underline{q})],$$

subject to

$$\begin{aligned}
\overline{\theta}\overline{q} - \overline{t} &\geq \overline{\theta}\underline{q} - \underline{t} && (\overline{IC}),\\
\underline{\theta}\underline{q} - \underline{t} &\geq \underline{\theta}\overline{q} - \overline{t} && (\underline{IC}),\\
\overline{\theta}\overline{q} - \overline{t} &\geq 0 && (\overline{IR}),\\
\underline{\theta}\underline{q} - \underline{t} &\geq 0 && (\underline{IR}),
\end{aligned}$$

where (IC) refers to an incentive compatibility constraint to choose the relevant contract and (IR) refers to an individual rationality constraint to choose some contract rather than no purchase at all.

Note that the two IC constraints can be combined in the statement

$$\overline{\theta}\Delta q \geq \Delta t \geq \underline{\theta}\Delta q.$$

Among other things we find that incentive compatibility implies that $\overline{q} \geq \underline{q}$.

To simplify the maximization program facing the firm, we consider the four constraints to determine which – if any – will be binding.

1. First note that $\overline{IC}$ and $\underline{IR}$ imply that $\overline{IR}$ is slack. Hence, it will never be binding and we can ignore it.

2. A simple argument establishes that $\overline{IC}$ must always bind. Suppose otherwise and it was slack at the optimal contract offering. In such a case, $\overline{\theta}$ could be raised slightly without disturbing this constraint or $\overline{IR}$, thereby increasing profits. Moreover, this increase only eases the $\underline{IC}$ constraint. Hence, the contract cannot be optimal. $\overline{IC}$ binds.

3. If $\overline{IC}$ binds, $\underline{IC}$ must be slack if $\overline{q} - \underline{q} \geq 0$ because

$$\overline{\theta}\Delta q = \Delta t > \underline{\theta}\Delta q.$$

Hence, we can ignore $\underline{IC}$ if we assume $\overline{q} - \underline{q} \geq 0$.

Because we have two constraints satisfied with equalities, we can use them to solve for $\overline{t}$ and $\underline{t}$ as functions of $\overline{q}$ and $\underline{q}$:

$$\begin{aligned}\underline{t} &= \underline{\theta}\underline{q}, \\ \overline{t} &= \underline{t} + \overline{\theta}\Delta q \\ &= \overline{\theta}\overline{q} - \Delta\theta\underline{q}.\end{aligned}$$

These t's are necessary and sufficient for all four constraints to be satisfied if $\overline{q} - \underline{q} \geq 0$. Substituting for the t's in the firm's objective program, the firm's maximization program becomes simply

$$\max_{\{(\underline{q},\underline{t}),(\overline{q},\overline{t})\}} p[\overline{\theta}\overline{q} - c(\overline{q}) - \Delta\underline{q}] + (1-p)[\underline{\theta}\underline{q} - c(\underline{q})],$$

subject to $\overline{q} \geq \underline{q}$. Ignoring the monotonicity constraint, the first-order conditions to this relaxed program imply

$$\overline{\theta} = c'(\overline{q}),$$

$$\underline{\theta} = c'(\underline{q}) + \frac{p}{1-p}\Delta\theta.$$

Hence, $\overline{q}$ is set at the first-best efficient levels of consumption but $\underline{q}$ is set at sub-optimal levels of consumption. This distortion also implies that $\overline{q} > \underline{q}$, and hence our monotonicity constraint does not bind.

Having determined the optimal $\overline{q}$ and $\underline{q}$, the firm can easily determine the appropriate prices using the conditions for $\overline{t}$ and $\underline{t}$ above and the firm's nonlinear pricing problem has been solved.

3.2.2 The Basic Paradigm with a Continuum of Types

This subsection borrows from Fudenberg-Tirole [Ch. 7, 1991], although many of the assumptions, theorems, and proofs have been modified.

There are usually two steps in mechanism design. First, characterizing the set of implementable contracts; then, selecting the optimal contract from this set.

First, some notation. For now, we consider the simpler case of a single agent, and so we have dropped subscripts. Additionally, it does not matter whether we focus on BNE or DSE allocations. The basic elements of the simple model are as follows:

1. Our allocation is a pair of non-stochastic functions, $y = (x, t)$, where $x \in \mathbb{R}_+$ is a one-dimensional activity (e.g., consumption, production, etc.) and t is a transfer (perhaps negative) from the agent to the principal. We will sometimes refer to x as the decision or activity of the agent and t as the transfer function.

2. The agent's private information is one-dimensional, $\theta \in \Theta$, where we take $\Theta = [0, 1]$ without loss of generality. The density is $p(\theta) > 0$ over $[0, 1]$, and the distribution function is $P(\theta)$.

3. The agent has quasi-linear utility: $U = u(x, \theta) - t$, where $u \in C^2$.

4. The principal has quasi-linear utility: $V = v(x, \theta) + t$, where $v \in C^2$.

5. The total surplus function is $S(x,\theta) \equiv u(x,\theta) + v(x,\theta)$; we assume that utilities are transferable.

Implementable Contracts Unlike F&T, we begin by making the Spence-Mirrlees single-crossing property (sorting) assumption on u.

Assumption 1 $\frac{\partial u(x,\theta)}{\partial \theta} > 0$ and $\frac{\partial^2 u(x,\theta)}{\partial \theta \partial x} > 0$.

The first condition is not generally considered part of the sorting condition, but because they are so closely related economically, we make them together.

Definition 6 We say that an allocation $y = (x,t)$ is **implementable** (or alternatively, we say that x is **implementable** with transfer t) iff it satisfies the incentive-compatibility (truth-telling) constraint

$$u(x(\theta),\theta) - t(\theta) \geq u(x(\hat{\theta}),\theta) - t(\hat{\theta}), \text{ for all } (\theta,\hat{\theta}) \in [0,1]^2. \tag{IC}$$

For notational ease, we will find it useful to consider the indirect utility function; i.e. the utility the agent of type θ receives when reporting $\hat{\theta}$: $U(\hat{\theta}|\theta) \equiv u(x(\hat{\theta}),\theta) - t(\hat{\theta})$. We will use the subscripts 1 and 2 to represent the partial derivatives of U with respect to report and type, respectively. When evaluating U in truthtelling equilibrium, we will often write $U(\theta) \equiv U(\theta|\theta)$. Note that in this case, $\frac{dU(\theta)}{d\theta} = U_1(\theta) + U_2(\theta)$. Our characterization theorem can now be presented and proved.

Theorem: Suppose $u_{x\theta} > 0$ and that the direct mechanism, $y(\theta) = (x(\theta), t(\theta))$, is is compact-valued (i.e., the set $\{(x,t) | \exists \hat{\theta} \in \Theta \text{ s.t. } (x,t) = (x(\hat{\theta}), t(\hat{\theta}))\}$ is compact). Then the direct mechanism is incentive compatible iff

$$U(\theta_1) - U(\theta_0) = \int_{\theta_0}^{\theta_1} u_\theta(x(s),s)ds, \quad \forall \theta_0, \theta_1 \in \Theta, \tag{3}$$

and $x(\theta)$ is nondecreasing.

The result in equation (3) is a restatement of the agent's first-order condition for truth-telling. Providing the mechanism is differentiable, when truth-telling is optimal we have $U_1(\theta) = 0$, and so $\frac{dU(\theta)}{d\theta} = U_2(\theta)$. Because $U_2(\theta) = u_\theta(x,\theta)$, applying the fundamental theorem of calculus yields equation (3). As the proof below makes clear, the monotonicity condition is the analog of the agent's second-order condition for truth-telling.

Proof:
Necessity: Incentive compatibility requires for any θ and $\tilde{\theta}$,

$$U(\theta) \geq U(\tilde{\theta}|\theta) \equiv U(\tilde{\theta}) + [u(x(\tilde{\theta}),\theta) - u(x(\tilde{\theta}),\tilde{\theta})].$$

Thus,

$$U(\theta) - U(\tilde{\theta}) \geq u(x(\tilde{\theta}),\theta) - u(x(\tilde{\theta}),\tilde{\theta}).$$

Reversing the θ and $\tilde{\theta}$ and combining results yields

$$u(x(\theta),\theta) - u(x(\theta),\tilde{\theta}) \geq U(\theta) - U(\tilde{\theta}) \geq u(x(\tilde{\theta}),\theta) - u(x(\tilde{\theta}),\tilde{\theta}).$$

Monotonicity is immediate from $u_{x\theta} > 0$.

Taking the limit as $\theta \to \hat{\theta}$ implies that

$$\frac{dU(\theta)}{d\theta} = u_\theta(x(\theta), \theta)$$

at all points at which $x(\theta)$ is continuous (which is everywhere but perhaps a countable number of points due to the monotonicity of x). Given that the set of available allocations is compact, continuity of $u(x, \theta)$ implies that $U(\theta)$ is continuous (by the Maximum theorem of Berges (1953)). The continuity of $U(\theta)$ over the compact set Θ (which implies U is uniformly continuous), combined with a bounded derivative (at all points of existence), implies that the fundamental theorem of calculus can be applied (specifically, U is Lipschitz continuous, and therefore absolutely continuous). Hence, $U(\theta)$ can be represented as in (3).

Sufficiency: Suppose not. Then there exists θ and $\hat{\theta}$ such that

$$U(\hat{\theta}|\theta) > U(\theta|\theta),$$

which implies

$$u(x(\hat{\theta}), \theta) - u(x(\hat{\theta}), \hat{\theta}) > U(\theta) - U(\hat{\theta}).$$

Integrating the lefthandside and using (1) above on the righthand side implies

$$\int_{\hat{\theta}}^{\theta} u_\theta(x(\hat{\theta}), s)ds > \int_{\hat{\theta}}^{\theta} u_\theta(x(s), s)ds.$$

Rearranging,

$$\int_{\hat{\theta}}^{\theta} [u_\theta(x(\hat{\theta}), s) - u_\theta(x(s), s)]ds > 0.$$

But the single-crossing property of A.1 that $u_{x\theta} > 0$ with the monotonicity condition implies that this is not possible. Hence, a contradiction.□

Remarks:

1. The above characterization theorem was first used by Mirrlees [1971] in his study of optimal taxation. Needless to say, it is a very powerful and useful theorem.

2. We could have used an alternative representation of the single-crossing property where type has a negative interpretation: $u_\theta < 0$ and $u_{x\theta} < 0$. This is isomorphic to our original sorting condition where θ is replaced by $-\theta$. Consequently, our characterization theorem is unchanged except that x must be nonincreasing. This alternative representation is commonly used in public regulation contexts where the type of the agent is related to marginal cost of production, and so higher types have higher costs and lower payoffs.

3. The implementation theorem above is actually easily generalized to cases of non-quasi-linear utility. In such a case, we need to make a Lipschitz assumption on the marginal rate of substitution between money and decision in order to guarantee the existence of a transfer function which can implement the decision function x. See Guesnerie and Laffont [1984] for details.

4. Related characterization theorems have been proved for the case of multi-dimensional types, multi-dimensional actions, multiple mechanisms (common agency). They all proceed in basically the same way, but rely on gradients and Hessians instead of simple one-dimensional first and second-order conditions.

5. The above characterization theorem can also be easily extended to random allocation functions, $\tilde{y} = (\tilde{x}, \tilde{t})$ by taking the appropriate expectations in the initial definition of the indirect utility function.

6. Unlike many papers in the literature, this statement uses the intergral condition (rather than the derivative condition) as the first-order condition. The reason for this is two-fold. First, the integral condition is what we ultimately would like to use for sufficiency (i.e., it is the fundamental theorem of calculus), and second, if the mechanism we are intersted in is not differentiable, the derivative condition is less useful. The difficulty over traditional proofs is then to show that the integral condition in (1) is actually a necessary condition (rather than the possibly weaker derivative condition). This is Myerson's approach in his proof in the "optimal auctions" paper. Myerson, however, leaves out the technical details in proving the necessity of (1) (i.e., that U can be integrated up to yield (3), which is more than saying that simply that U can be integrated), and in any event Myerson has the advantage of a simpler problem in that $u = \theta x$ in his framework; we do not have such a luxury. Note that to accomplish our result, I have used a requirement that the direct mechanism is compact-valued. Without such an assumption, a direct mechanism may not have an optimal report (technically, $\sup_{\hat{\theta}} U(\hat{\theta}|\theta)$ exists, but $\max_{\hat{\theta}} U(\hat{\theta}|\theta)$ may not). This seems a very unrestrictive notion to place on our contract space.

7. To reiterate, this stronger theorem for IC (without relying on continuity, etc.) is essential when looking at optimal auctions. The implementation theorem in Fudenberg and Tirole's book (ch. 7), for example, is not directly useful because they have resorted to assuming x is continuous and has first derivatives that are continuous at all but a finite number of points – far too stringent of a condition for auctions.

Optimal Contracts We now consider the optimal choice of contract by a principal. Given our characterization theorem, the principal's problem is

$$\max_{y} \; E_\theta[V(x(\theta), \theta)] \equiv E_\theta[S(x(\theta), \theta) - U(x(\theta), \theta)]$$

subject to $\frac{dU(\theta)}{d\theta} = u_\theta(x(\theta), \theta)$, x nondecreasing, and generally a participation constraint (referred to as individual rationality in the literature)

$$U(\theta) \geq \underline{U}, \text{ for all } \theta \in [0,1]. \tag{IR}$$

[Note that we have rewritten profits as the difference between total surplus and the agent's surplus.]

Normally, one proceeds by solving the relaxed program in which monotonicity is ignored, and then checking ex post that it is in fact satisfied. If it isn't satisfied, one then either incorporates the constraint into the maximization program directly (a tedious thing to do) or assume sufficient regularity conditions so that the resulting function is indeed monotonic. We will also follow this latter approach for now.

In solving the relaxed program, there are again two choices available. First, and when possible I think the most powerful, integrating out the agent's utility function and converting the problem to one of pointwise maximization; second, using control theoretic tools (e.g., Hamiltonians and Pontryagin's theorem) directly to solve the problem. We will begin with the former.

Note that the second term of the objective function can be rewritten using integration by parts and the fact that $\frac{dU}{d\theta} = u_\theta$.

$$\begin{aligned} E_\theta[U(x(\theta),\theta)] &\equiv \int_0^1 U(x(s),s)p(s)ds \\ &= -U(x(\theta),\theta)[1-P(\theta)]|_0^1 + \int_0^1 \frac{dU(s)}{d\theta}\frac{1-P(s)}{p(s)}p(s)ds \\ &= U(x(0),0) + E_\theta\left[u_\theta(x(\theta),\theta)\frac{1-P(\theta)}{p(\theta)}\right]. \end{aligned}$$

Remark: It is equally true (via changing the constant of integration) that

$$E_\theta[U(x(\theta),\theta)] = U(x(1),1) - E_\theta\left[u_\theta(x(\theta),\theta)\frac{P(\theta)}{p(\theta)}\right].$$

We use the former representation rather than the latter because when $u_\theta > 0$, it will typically be optimal to set $U(x(0),0)$ equal to the agent's outside reservation utility, $\underline{U}$; $U(x(1),1)$ on the other hand is endogenously determined. This is true since utility is increasing in type, so the participation constraint will bind only for the lowest type. When the alternative sorting condition is in place (i.e., $u_\theta < 0$ and $u_{x\theta} < 0$), the second representation will be more useful as $U(x(1),1)$ will typically be set to the agent's outside utility.

Now we substitute the new representation of the agent's expected utility for the one in the objective function. We have our relaxed program (ignoring the monotonicity constraint) as

$$\max_x E_\theta\left[S(x(\theta),\theta) - \frac{1-P(\theta)}{p(\theta)}u_\theta(x(\theta),\theta) - U(x(0),0)\right],$$

subject to (IR) and (3). For notational ease, define

$$\Phi(x,\theta) \equiv S(x,\theta) - \frac{1-P(\theta)}{p(\theta)}u_\theta(x,\theta).$$

We make the following regularity assumption.

Assumption 2 *Φ is quasi-concave and has a unique interior maximum over $x \in \mathbb{R}_+$ $\forall\theta$.*

This assumption is uncontroversial and is met, for example, if the underlying surplus function is strictly concave, $-u_\theta$ is not too convex, and $\Phi_x(0,\theta)$ is nonnegative. We now have a theorem which characterizes the solution to our relaxed program.

Theorem 21 Suppose that A.2 holds, that x satisfies $\Phi_x(x(\theta),\theta) = 0$ $\forall\theta \in [0,1]$, and that $t(\theta) = u(x(\theta),\theta) - \left(\underline{U} + \int_0^\theta u_\theta(x(s),s)ds\right)$. Then $y = (x,t)$ solves the relaxed program.

Proof: The proof follows immediately from our simplified objective function. First, note that the objective function has been re-written independent of transfers. [When we integrated by parts, we integrated out the transfer function because U is quasi-linear in t.] Thus, we can choose x to maximize the objective function and later choose t such that the differential equation (3) is satisfied for that x; i.e., $t = u - U$, or

$$t(\theta) = u(x(\theta), \theta) - \left(U(x(0), 0) + \int_0^\theta u_\theta(x(s), s) ds \right).$$

Given that $\frac{dU}{d\theta} > 0$, the IR constraint can be restated as $U(x(0), 0) \geq \underline{U}$. From the objective function, this constraint will bind because there is never any reason to leave the lowest type rents. Hence, we have our equation for transfers given x. Finally, to obtain the optimal x, note that our choice of x solves $\max_x E[\Phi(x(\theta), \theta)]$. The optimal x will maximize $\Phi(x, \theta)$ pointwise in θ, which by A.2, is equivalent to $\Phi_x(x(\theta), \theta) = 0 \; \forall \theta$. □

We still have to check that x is nondecreasing in order for $y = (x, t)$ to be an optimal contract. The necessary and sufficient condition for this to be true is given in the following regularity condition.

Assumption 3 $\Phi_{x\theta} \geq 0$ for all (x, θ).

Remark: This assumption is not standard, but is much weaker and more intuitive than that commonly made in the literature. Typically, the literature assumes that $v_{x\theta} \geq 0$, $u_{x\theta\theta} \leq 0$, and the distribution of types satisfies a monotone hazard-rate condition (MHRC). These three conditions imply A.3. The first two assumptions are straightforward, although we generally have little economic insight into third derivatives. The hazard rate is $\frac{p}{1-P}$, and the MHRC assumes that this is nondecreasing. This assumption is satisfied for several common distributions such as the uniform, normal, logistic, and exponential, for example.

Theorem 22 Suppose that A.2 and A.3 are satisfied. Then the solution to the relaxed program satisfies the original un-relaxed program.

Proof: Differentiating $\Phi_x(x(\theta), \theta) = 0$ with respect to θ implies $\frac{dx(\theta)}{d\theta} = -\frac{\Phi_{x\theta}}{\Phi_{xx}}$. By A.2, the denominator is negative. By A.3, x is nondecreasing. □

Given our regularity conditions, we know that the optimal contract satisfies

$$\Phi_x(x(\theta), \theta) = 0.$$

What does this mean?

Interpretations of $\Phi_x(x(\theta), \theta) = 0$:

1. We can rewrite the optimality condition for x as

$$S_x(x(\theta), \theta) = \frac{1 - P(\theta)}{p(\theta)} u_{x\theta}(x(\theta), \theta) \geq 0.$$

   Clearly, there is an under-provision of the contracted activity, x, for all but the highest type. For the highest type, $\theta = 1$, we have the full-information level of activity.

2. Alternatively, we can rewrite the optimality condition for x as

$$p(\theta)S_x(x(\theta),\theta) = [1 - P(\theta)]u_{x\theta}(x(\theta),\theta).$$

Fix a particular value of θ. The LHS represents the marginal gain in joint surplus by increasing $x(\theta)$ to $x(\theta) + dx$. It is multiplied by $p(\theta)$ which represents the probability of a type occurring between θ and $\theta + d\theta$. The RHS represents the marginal cost of increasing x at th: for all higher types, rents will be increased. Remember that the agent's rent increases by u_θ in θ. Because u_θ increases in x, a small increase in x implies that rents will be increased by $u_{x\theta}$ for all types above θ, which exist with probability $1 - P(\theta)$.

It is very similar to the distortion a monopolist introduces to maximize profits. Let the buyer's unit value of a good be distributed according to $F(v)$. The buyer buys a unit iff the value is greater than price, p; marginal cost is constant at c. Thus, the monopolist solves $\max_p [1 - F(p)](p - c)$, which yields as a first-order condition,

$$f(p)(p - c) = [1 - F(p)].$$

Lowering the price increases profits on the marginal consumer (LHS) but lowers profits on all inframarginal customers who would have purchased at the higher price (RHS).

3. Following Myerson [1981], we can redefine the agent's utility as a *virtual utility* which represents the agent's utility less information rents:

$$\tilde{u}(x,\theta) \equiv u(x,\theta) - \frac{1 - P(\theta)}{p(\theta)} u_\theta(x,\theta).$$

The principal maximizes the sum of the virtual utilities. This terminology is particular useful in the study of optimal auctions.

Remarks on Basic Paradigm:

1. If it is unreasonable to assume A.3 for the economic problem under study, one must maximize the un-relaxed program including a constraint for the monotonicity condition. The technique is straightforward, but tedious. It involves optimally "ironing out" the decision function found in the relaxed program. That is, carefully choosing intervals where x is made constant, but otherwise following the relaxed choice of x. The result is that there will be regions of pooling in the optimal contract, but in non-pooling regions, it will be the same as before. Additionally, there will typically (although not always) be no distortion at the top and insufficient x for all other types.

2. A few brave researchers have extended the optimality results to cases where there are several dimensions of private information. Roughly, the trick is to note that a multidimensional incentive problem can be converted to a single-dimensional one by defining a new type using the agent's indifference curves over the old type space. Mathematically, rather than integrating by parts, Stoke's theorem can be used. See the cites in F&T for more information, or look at Wilson's [1993] book, *Nonlinear Pricing*, or at the papers by Armstrong [1996] and Rochet [1995].

3. Risk aversion (i.e., non-quasi-linear preferences) does not affect the implementability theorem significantly, but does affect the choice of contract. Importantly, wealth effects may alter the optimal contract dramatically. Salanie [1990] and Laffont and Rochet [1994] consider these problems in detail finding that a region of pooling at the low-end of the distribution occurs for some intermediate levels of risk aversion.

4. Common agency. With two or more principals, there will be externalities in contract choice. Just like two duopolists erode some of the available monopoly profits, so will two principals. What's interesting is the conduit for the erosion. Simply stated, principals competing over substitute (complementary) activities will reduce (increase) the distortion the agent faces. See Martimort [1992,1996] and Stole [1990] for details.

5. We've assumed that the agent knows more than the principal. But suppose that its the other way around. Now it is possible that the principal will signal her private information in the contract offer. Thus, we are looking at signaling contracts rather than screening contracts. The results are useful for our purposes, however, as sometimes we may want to consider renegotiation led by a privately informed agent. We'll talk more about these issues later in the course. The relevant papers are Maskin and Tirole [1990,1992].

6. We limited our attention to deterministic mechanisms when searching for the optimal mechanism. Could stochastic mechanisms do any better? If the surplus function is concave in x, the only possible value is that a stochastic mechanism might reduce the rent term of the agent; i.e., u_θ may be concave. Most researchers assume that $u_{\theta xx} \geq 0$ which is sufficient to rule out stochastic mechanisms.

7. There is a completely different approach to optimal contract design which focuses on probability distributions over cumulative economic activity at given tariffs rather than distributions over types. This approach was first used by Goldman-Leland-Sibley [1984] and has recently been put to great use in Wilson's *Nonlinear Pricing*. Of course, it yields identical answers, but with very different mathematics. My sense is that this framework may give testable implications for demand data more directly than using the approach developed above.

8. We have derived the optimal direct revelation mechanism contract. Are there economically reasonable indirect mechanisms which yield the same allocations and that we expect to see played? Two come to mind. First, because x and t are monotonic, we can construct a nonlinear tariff, $T(x)$, which implements the same allocation; here, the agent is allowed to choose from the menu. Second, if $T(x)$ is concave (which we will see occurs frequently), an optimal indirect mechanism of the form of a menu of linear contracts exists, where the agent chooses a particular two-part tariff and can consume anywhere along it. This type of contract has a particularly nice robustness against noise, which we will see when we study Laffont-Tirole [1986], where the idea of a menu of two-part tariff was developed.

9. Many researchers use control theory to solve the problem rather than integration by parts and pointwise maximization. The cost of this approach is that the standard sufficient conditions for an optimal solution in control theory are stronger than what

we used above. Additionally, a Hamiltonian has no immediately clear economic interpretation. The benefits are that sometimes the tricks used above cannot be used. Control theory is far more powerful and general than our simple integration by parts trick. So for complicated problems, it is something which can be useful. The basic idea is to treat θ as the state variable just as engineers would treat time. The control variable is x and the co-state variable is indirect utility, U. The Hamiltonian becomes

$$\mathcal{H}(x, U, \theta) \equiv (S(x, \theta) - U)p(\theta) + \lambda(\theta)u_\theta(x, \theta).$$

Roughly speaking, providing x is piecewise-C^1 and $\mathcal{H}$ is globally strictly concave in (x, U) for any λ (this puts lots of restrictions on u and v), the following conditions are necessary and sufficient for an optimum:

$$\mathcal{H}_x(x(\theta), U(\theta), \theta) = 0,$$

$$-\mathcal{H}_U(x, U, \theta) = \lambda'(\theta),$$

$$\lambda(1) = 0.$$

Solving these equations yields the same solution as above. Weaker versions of the concavity conditions are available; see for example Seierstad and Sydsaeter's [1987] control theory book for details.

3.2.3 Finite Distribution of Types

Rather than use continuous distributions of types and deal with the functional analysis messes (Lipschitz conditions, etc.), some of the literature has used finite distributions. Most notable are Hart [1983] and Moore [1988]. The approach (first characterize implementable contracts, then optimize) is the same. The techniques of the proofs are different enough to warrant some attention.

For the purposes of this section, suppose that finite versions of A.2-A.3 still hold, but that there are only n types, $\theta_1 < \theta_2, \ldots, \theta_{n-1} < \theta_n$ with probability "density" p_i for each type and "distribution" function $P_i \equiv \sum_{j=1}^{i} p_j$. Thus, $P_1 = p_1$ and $P_n = 1$. Let $y_i = (x_i, t_i)$ represent the allocation to the agent claiming to be the ith type.

Implementable Contracts Our principal's program is to

$$\max_{y_i} \sum_{i=1}^{n} p_i \left\{ S(x_i, \theta_i) - U(\theta_i) \right\},$$

subject to, $\forall\, i, j,$,

$$U(\theta_i|\theta_i) \geq U(\theta_j|\theta_i), \tag{IC(i,j)}$$

and

$$U(\theta_i|\theta_i) \geq \underline{U}. \tag{IR(i)}$$

Generally, we can eliminate many of these constraints and focus on local incentive compatibility.

Theorem 23 If $u_{x\theta} \geq 0$, then the local constraints

$$U(\theta_i|\theta_i) \geq U(\theta_{i-1}|\theta_i) \qquad \text{(DLIC(i))}$$

and

$$U(\theta_i|\theta_i) \geq U(\theta_{i+1}|\theta_i) \qquad \text{(ULIC(i))}$$

satisfied for all i are necessary and sufficient for global incentive compatibility.

Proof: Necessity is direct. Sufficiency is proven by induction.

First, note that the local constraints imply that $x_i \geq x_{i-1}$. Specifically, the local constraints imply

$$U(\theta_i|\theta_i) - U(\theta_{i-1}|\theta_i) \geq 0 \geq U(\theta_i|\theta_{i-1}) - U(\theta_{i-1}|\theta_{i-1}), \;\; \forall\, i.$$

Rearranging, and using direct utilities, we have

$$u(x_i, \theta_i) - u(x_{i-1}, \theta_i) \geq u(x_i, \theta_{i-1}) - u(x_{i-1}, \theta_{i-1}).$$

Combining this inequality with the sorting condition implies monotonicity.

Consider DLIC for type i and $i-1$. Restated in direct utility terms, these conditions are

$$u(x_i, \theta_i) - u(x_{i-1}, \theta_i) \geq t_i - t_{i-1},$$
$$u(x_{i-1}, \theta_{i-1}) - u(x_{i-2}, \theta_{i-1}) \geq t_{i-1} - t_{i-2}.$$

Adding the conditions imply

$$u(x_i, \theta_i) - u(x_{i-1}, \theta_i) + u(x_{i-1}, \theta_{i-1}) - u(x_{i-2}, \theta_{i-1}) \geq t_i - t_{i-2}.$$

By the sorting condition and monotonicity, the LHS is smaller than

$$u(x_i, \theta_i) - u(x_{i-1}, \theta_i) + u(x_{i-1}, \theta_i) - u(x_{i-2}, \theta_i) = u(x_i, \theta_i) - u(x_{i-2}, \theta_i),$$

and so IC(i,i-2) is satisfied:

$$u(x_i, \theta_i) - u(x_{i-2}, \theta_i) \geq t_i - t_{i-2}.$$

Thus, DLIC(i) and DLIC(i-1) imply IC(i,i-2). One can show that IC(i,i-1) and DLIC(i-2) imply IC(i,i-3), etc. Therefore, starting at $i = n$ and proceeding inductively, DLIC implies IC(i,j) holds for all $i \geq j$. A similar argument in the reverse direction establishes that ULIC implies IC(i,j) for $i \leq j$. □

The basic idea of the theorem is that the local upward constraints imply global upward constraints, and likewise for the downward constraints. We have reduced our necessary and sufficient IC constraints from $n(n-1)$ to $2(n-1)$ constraints. We can now optimize using Kuhn-Tucker's theorem. If possible, however, it is better to check to see if we can simplify things a bit more.

We can still do better for our *particular* problem. Consider the following **relaxed program**.

$$\max_{y_i} \sum_{i=1}^{n} p_i \left\{ S(x_i, \theta_i) - U(\theta_i) \right\},$$

subject to DLIC(i) for every i, IR(1), and x_i nondecreasing in i.

We will demonstrate that

Theorem 24 The solution to the unrelaxed program is equivalent to the solution of the relaxed program.

Proof: The proof proceeds in 3 steps.

Step 1: The constraints of the unrelaxed program imply those of the relaxed program. It is easy to see that IC(i,j) imply DLIC(i) and IR(i) imply IR(1). Take $i > j$. By IC(i,j) and IC(j,i) we have

$$u(x_i, \theta_i) - t_i \geq u(x_j, \theta_i) - t_j,$$
$$u(x_j, \theta_j) - t_j \geq u(x_i, \theta_j) - t_i.$$

Adding and rearranging,

$$[u(x_i, \theta_i) - u(x_j, \theta_i)] - [u(x_i, \theta_j) - u(x_j, \theta_j)] \geq 0.$$

By the sorting condition, if $\theta_i > \theta_j$, then $x_i \geq x_j$.

Step 2: At the solution of the relaxed program, DLIC(i) is binding for all i. Suppose not. Take i and ε such that $[u(x_i, \theta_i) - t_i] - [u(x_{i-1}, \theta_i) - t_{i-1}] > \varepsilon > 0$. Now for all $j \geq i$, raise transfers to $t_j + \varepsilon$. No IC constraints will be violated and profit is raised by $(1 - P_{i-1})\varepsilon$, which contradicts $\{x, t\}$ being a solution to the relaxed program.

Step 3: The solution of the relaxed program satisfies the constraints of the unrelaxed program. Because DLIC(i) is binding, we have

$$u(x_i, \theta_i) - u(x_{i-1}, \theta_i) = t_i - t_{i-1}.$$

By monotonicity and the sorting condition,

$$u(x_i, \theta_{i-1}) - u(x_{i-1}, \theta_{i-1}) \leq t_i - t_{i-1}.$$

But this latter condition is ULIC(i-1). Hence, DLIC and ULIC are satisfied. By Theorem 23, this is sufficient for global incentive compatibility. Finally, it is straightforward to show that IC(i,j) and IR(1) implies IR(i). □

Optimal Contracts We now solve the simpler relaxed program. Note that there are now only n constraints, all of which are binding so we can use Lagrangian analysis rather than Kuhn-Tucker conditions given appropriate assumptions of concavity and convexity.

We solve

$$\max_{x_i, U_i} \mathcal{L} = \sum_{i=1}^{n} p_i \{S(x_i, \theta_i) - U_i\} + \sum_{i=2}^{n} \lambda_i (U_i - U_{i-1} - u(x_{i-1}, \theta_i) + u(x_{i-1}, \theta_{i-1})) + \lambda_1 (U_1 - \underline{U}),$$

ignoring the monotonicity constraint for now. We have used $U_i \equiv U(\theta_i)$ instead of transfers as our primitive instruments, along the lines of the control theory approach. [Using indirect utilities, the DLIC constraints become $U_i - U_{i-1} - u(x_{i-1}, \theta_i) + u(x_{i-1}, \theta_{i-1}) \geq 0$.] There are $2n$ necessary first-order conditions:

$$p_i S_x(x_i, \theta_i) = \lambda_{i+1}[u_x(x_i, \theta_{i+1}) - u_x(x_i, \theta_i)],\ i = 1, \ldots, n-1,$$

$$p_n S_x(x_n, \theta_n) = 0,$$

$$-p_i + \lambda_i - \lambda_{i+1} = 0,\ i = 1, \ldots, n-1,$$

$$-p_n + \lambda_n = 0.$$

Combining the last two sets of equations, we have a first-order difference equation which we can solve uniquely for $\lambda_i = \sum_{j=i}^{n} p_j$. Thus, assuming discrete analogues of A.2 and A.3, we have the following result.

Theorem 25 In the case of a finite distribution of types, the optimal mechanism has x_i to satisfy

$$p_i S_x(x_i, \theta_i) = [1 - P_i](u_x(x_i, \theta_{i+1}) - u_x(x_i, \theta_i)),\ i = 1, ..., n,$$

t_i is chosen as the unique solution to the first-order difference equation,

$$U_i - U_{i-1} = u(x_{i-1}, \theta_i) + u(x_{i-1}, \theta_{i-1}),$$

with initial condition, $U_1 = \underline{U}$.

Remarks:

1. As before, we have no distortion at the top and a suboptimal level of activity for all lower types.

2. Step 2 of the proof of Theorem 24 is frequently skipped by researchers. Be careful because DLIC does not bind in all economic environments. Moreover, it is incorrect to assume DLIC binds (i.e., impose DLIC's with equalities in the relaxed program) and then check that the other constraints are satisfied in the solution to the relaxed program. This does not guarantee an optimum!!! Either you must show that the constraints are binding or you must use Kuhn-Tucker analysis which allows the constraints to bind or be slack. This is important.

   In many papers, it is not the case that the DLIC constraints bind. See, Hart, [1983] for example. In fact, in Stole [1996], there is an example of a price discrimination problem where the upward local IC constraints bind rather than the downward ones. This is generated by adding noise to the reservation utility of consumers. As a consequence, you may want to leave rents to some types to increase the chances that they will visit your store, but then the logic of step 2 does not work and in fact the upward constraints may bind.

3. Discrete models sometimes generate results which are quite different from those which emerge from the continuous-type settings. For example, in Stole [1996], the discrete setting with random IR constraints exhibits a downward distortion for the lowest type which is absent in the continuous setting. As another example, provided by David Martimort, in a Riley and Samuelson [1981] auction setting with a continuum of types and the seller's reservation price set equal to the buyer's lowest valuation, there is a multiplicity of symmetric equilibria (this is due to a singularity in the characterizing differential equation at $\theta = \underline{\theta}$). With discrete types, a unique symmetric equilibrium exists.

4. It is frequently convenient to use two-type models to get the feel for an economic problem before going on to tackle the n-type or continuous-type case. While this is usually helpful, some economic phenomena will only be visible in $n \geq 3$ environments. (E.g., common agency models with complements generate the first-best as an outcome with two types but not with three or more.) Fundamentally, the IC constraints in the discrete setting are much more rigid with two types than with a continuum. Nonetheless, much can be learned from two-type models, although it would certainly be better if the result could be generalized to more.

3.2.4 Application: Nonlinear Pricing

We now turn to the case of nonlinear pricing developed by Goldman, Leland and Sibley [1984], Mussa-Rosen [1978], and Maskin-Riley [1984]. Given the development for the single-agent case above, this is a simple exercise of applying our theorems.

Basic Model: Consider a monopolist who faces a population of customers with varying marginal valuations for its product. Let the consumers be indexed by type, $\theta \in [0, 1]$, with utility functions $U = u(x, \theta) - t$. We will assume that a higher type customer receives both greater total and greater marginal utility from consumption, so A.1 is satisfied. The monopolist's payoffs for a given sale of x units is $V = t - C(x)$. The monopolist wishes to find the optimal nonlinear price schedule to offer its customers.

Note that $\Phi(x, \theta) \equiv u(x, \theta) - C(x) - \frac{1-P(\theta)}{p(\theta)} u_\theta(x, \theta)$ in this setting. We further assume that A.2 and A.3 are satisfied for this function.

Results:

1. Theorems 3.2.2, 21 and 22 imply that the optimal nonlinear contract satisfies

$$p(\theta)(u_x(x, \theta) - C_x(x)) = [1 - P(\theta)]u_{x\theta}(x, \theta).$$

Thus we have the result that quantity is optimally provided for the highest type and under-provided for lower types.

2. If we are prepared to assume a simpler functional form for utility, we can simplify this expression further. Let $u(x, \theta) \equiv \theta\nu(x)$. Then,

$$p(\theta)(\theta\nu_x(x) - C_x(x)) = [1 - P(\theta)]\nu_x(x).$$

Note that the marginal price a consumer of type θ pays for the marginal unit purchased is $T_x(x(\theta)) \equiv t'(\theta)/x'(\theta) = u_x(x(\theta),\theta)$. Rearranging our optimality condition, we have

$$\frac{T_x - C_x}{T_x} = \frac{1 - P(\theta)}{\theta p(\theta)}.$$

If P satisfies the monotone hazard-rate condition, then the Lerner index at the optimal allocation is decreasing in type. Note also that the RHS can be reinterpreted as the inverse virtual elasticity of demand.

3. Returning to our general model, $U = u(x,\theta) - t$, if we assume that marginal costs are constant, MHRC holds, $u_{x\theta\theta} \leq 0$, and $u_{xx\theta} \leq 0$, we can show that quantity discounts are optimal. [Note, all of these conditions are implied by the simple model above in result 2.] To show this, redefine the marginal nonlinear tariff schedule as $T_x(x) = u_x(x, x^{-1}(x))$, where $x^{-1}(x)$ gives the type θ which is willing to buy exactly x units; note that $T_x(x)$ is independent of θ. We want to show that $T(x)$ is a strictly concave function. Differentiating, $T_{xx} < 0$ is equivalent to

$$\frac{dx}{d\theta} > -\frac{u_{x\theta}(x,\theta)}{u_{xx}(x,\theta)}.$$

Because $x'(\theta) = -\Phi_{x\theta}/\Phi_{xx}$, the condition becomes

$$\frac{dx}{d\theta} = \frac{u_{x\theta} - \frac{1-P}{p} u_{x\theta\theta} - \left(\frac{d}{d\theta}\frac{1-P}{p}\right) u_{x\theta}}{-u_{xx} + \frac{1-P}{p} u_{x\theta\theta}} > -\frac{u_{x\theta}}{u_{xx}},$$

which is satisfied given our assumptions.

4. We could have instead written this model in terms of quality rather than quantity, where each consumer has unit demands but differing marginal valuations for quality. Nothing changes in the analysis. Just reinterpret x as quality.

3.2.5 Application: Regulation

The seminal papers in the theory of regulating a monopolist with unknown costs are Baron and Myerson [1982] and Laffont and Tirole [1986]. We will illustrate the main results of L&T here, and leave the results of B&M for a problem set or recitation.

Basic Model:

A regulated firm has private information about its costs, $\theta \in [\underline{\theta}, \overline{\theta}]$, distributed according to P, which we will assume satisfies the MHRC. The firm exerts an effort level, e, which has the effect of reducing the firm's marginal cost of production. The total cost of production is $C(q) = (\theta - e)q$. This effort, however, is costly; the firm's costs of effort are given by $\psi(e)$, which is increasing, strictly convex, and $\psi'''(e) \geq 0$. This last condition will imply that A.2 and A.3 are satisfied (as well as the optimality of non-random contracts). It is assumed that the regulator can observe costs, and so without loss of generality, we assume that the regulator pays the observed costs of production rather than the firm. Hence, the firm's utility is given by

$$U = t - \psi(e).$$

Laffont and Tirole assume a novel contracting structure in which the regulator can observe costs, but must determine how much of the costs are attributable to effort and how much are attributed to inherent luck (i.e., type). The regulator cannot observe e but can observe and contract upon total production costs C and output q. For any given q, the regulator can perfectly determine the firm's marginal cost $c = \theta - e$. Thus, the regulator can ask the firm to report its type and assign the firm a marginal cost target of $c(\hat{\theta})$ and an output level of $q(\hat{\theta})$ in exchange for compensation equal to $t(\hat{\theta})$. A firm with type θ that wishes to make the marginal cost target of $c(\hat{\theta})$ must expend effort equal to $e = \theta - c(\hat{\theta})$. With such a contract, the firm's indirect utility function becomes

$$U(\hat{\theta}|\theta) \equiv t(\hat{\theta}) - \psi(\theta - c(\hat{\theta})).$$

Note that this utility function is independent of $q(\hat{\theta})$ and that the sorting condition A.1 is satisfied if one normalizes θ to $-\theta$.

Theorem 3.2.2 implies that incentive compatible contracts are equivalent to requiring that

$$\frac{dU(\theta)}{d\theta} = -\psi'(\theta - c(\theta)),$$

and that $c(\theta)$ be nondecreasing.

To solve for the optimal contract, we need to state the regulator's objectives. Let's suppose that the regulator wishes to maximize a weighted average of strictly concave consumer surplus, $CS(q)$, (less costs and transfers) and producer surplus, U, with less weight afforded to the latter.

$$V = E_\theta[CS(q(\theta)) - c(\theta)q(\theta) - t(\theta) + \gamma U(\theta)],$$

or substituting out the transfer function,

$$V = E_\theta[CS(q(\theta)) - c(\theta)q(\theta) - \psi(\theta - c(\theta)) - (1-\gamma)U(\theta)],$$

where $0 \leq \gamma < 1$.

Remarks:

1. In our above development, we could instead write $S = CS - cq - \psi$, and then the regulator maximizes $E[S - (1-\gamma)U]$. If $\gamma = 0$, we are in the same situation as our initial framework where the principal doesn't directly value the agent's utility.

2. L&T motivate the cost of leaving rents to the firm as arising from the shadow cost of raising public funds. In that case, if $1 + \lambda$ is the cost of public funds, the regulator's objective function (after simplification) is $E[CS - (1+\lambda)(cq + \psi) - \lambda U]$. Except for the optimal choice of q, this yields identical results as using $1 - \gamma \equiv \frac{\lambda}{1+\lambda}$.

Our regulator solves the following program:

$$\max_{q,c,t} E_\theta[CS(q(\theta)) - c(\theta)q(\theta) - \psi(\theta - c(\theta)) - (1-\gamma)U(\theta)],$$

subject to

$$\frac{dU}{d\theta} = -\psi'(\theta - c(\theta)),$$

$c(\theta)$ nondecreasing, and the firm making nonnegative profits (i.e., $U(\theta) \geq 0$). Integrating $E_\theta[U(\theta)]$ by parts (using our previously developed techniques), we can substitute out U from the objective function (thereby eliminating transfers from the program). We then have

$$\max_{q,c} E_\theta \left[CS(q(\theta)) - c(\theta)q - \psi(\theta - c(\theta)) - (1-\gamma)\frac{P(\theta)}{p(\theta)}\psi'(\theta - c(\theta)) - (1-\gamma)U(\overline{\theta}) \right],$$

subject to transfers satisfying

$$\frac{dU}{d\theta} = -\psi'(\theta - c(\theta)),$$

$c(\theta)$ nondecreasing, and $U(\overline{\theta}) \geq 0$. We obtain the following results.

Results: Redefine for a given q

$$\Phi(c,\theta) \equiv CS(q) - cq - \psi(\theta - c) - (1-\gamma)\frac{P(\theta)}{p(\theta)}\psi'(\theta - c).$$

A.2 and A.3 are satisfied given our conditions on P and ψ so we can apply Theorems 21 and 22.

1. The choice of effort, $e(\theta)$, satisfies

   $$q(\theta) - \psi'(e(\theta)) = (1-\gamma)\frac{P(\theta)}{p(\theta)}\psi''(e(\theta)) \geq 0.$$

   Note that the first-best level of effort, conditional on q, is $q = \psi'(e)$. As a consequence, suboptimal effort is provided for every type except the lowest. We always have no distortion on the bottom. Only if $\gamma = 1$, so the regulator places equal weight on the firm's surplus, do we have no distortion anywhere. In the L&T setting, this condition translates to $\lambda = 0$; i.e., no excess cost of public funds.

2. The optimal q is the full-information efficient production level conditional on the marginal cost $c(\theta) \equiv \theta - e(\theta)$:

   $$CS'(q) = c(\theta).$$

   This is not the result of L&T. They find that because public funds are costly, $CS'(q) = (1+\lambda)c(\theta)$, where the RHS represents the effective marginal cost (taking into account public funds). Nonetheless, this solution still corresponds to the choice of q under full-information for a given marginal cost, because the cost of public funds is independent of informational issues. Hence, there is a dichotomy between pricing (choosing $c(\theta)$) and production.

3. Because $\psi''' \geq 0$, random schemes are never optimal.

4. L&T also show that the optimal nonlinear contract can be implemented using a realistic menu of two-part tariffs (i.e., cost-sharing contracts). Let $\overline{C}(q)$ be a total cost target. The firm chooses an output and a corresponding cost target, $\overline{C}(q)$, and then is compensated according to

   $$T(q,C) = F(q) + \alpha[\overline{C}(q) - C],$$

where C is observed ex post cost. F is interpreted as a fixed payment and α is a cost-sharing parameter. If $(\overline{\theta} - \underline{\theta})$ goes to zero, α goes to one, implying a fixed-price contract in which the firm absorbs any cost overruns. Of course, this framework doesn't make much sense with no other uncertainty in the model (i.e., there would never be cost overruns which we would observe). But if some noise is introduced on the ex post cost observation, i.e. $\tilde{C} = C(q) + \varepsilon$, the mechanism is still optimal. It is robust to linear noise in observed contract variables because the firm is risk neutral and the implemented contract is linear in the observation noise.

3.2.6 Resource Allocation Devices with Multiple Agents

We now consider problems with multiple agents, and so we will reintroduce our subscripts, $i = 1, \ldots, I$ to denote the different agents. We will also assume that the agent's types are independently distributed (but not necessarily identically), and so we will denote the individual density and distribution functions for $\theta_i \in [\underline{\theta}, \overline{\theta}]$ as $p_i(\theta_i)$ and $P_i(\theta_i)$, respectively. Also, we will sometimes use $p_{-i}(\theta_{-i}) \equiv \prod_{j \neq i} p_j(\theta_j)$.

Optimal Auctions This section closely follows Myerson [1981], albeit with different notation and some simplification on the type spaces.

We restrict attention to a single-unit auction in which the good is awarded to a one individual and consider Bayesian Nash implementation. There are I potential bidders for the object. The participants' expected utilities are given by

$$U_i = \phi_i \theta_i - t_i,$$

where ϕ_i is participant i's probability of receiving the good in the auction and θ_i is the marginal valuation for the good. t_i is the payment of the ith player to the principal (auctioneer). Note that because the agents are risk neutral, it is without loss of generality to consider payments made independently of getting the good. We will require that all bidders receive at least nonnegative payoffs: $U_i \geq 0$. This setup is known as the *Independent Private Values (IPV)* model of auctions. The "private" part refers to the fact that an individual's valuation is independent of what others think. As such, think of the object as a personal consumption good that will not be re-traded in the future.

The principal's direct mechanism is given by $y = (\phi, t)$. The agent's indirect utility functions given this mechanism can be stated as

$$U_i(\hat{\theta}|\theta_i) \equiv \phi_i(\hat{\theta})\theta_i - t_i(\hat{\theta}).$$

Note that the probability of winning and the transfer can depend upon everyone's announced type (as is typical in an auction setting). To simplify notation, we will indicate that the expectation of a function has been taken by eliminating the variables of integration from the function's argument. Specifically, define

$$\phi_i(\theta_i) \equiv E_{\theta_{-i}}[\phi_i(\theta)],$$

and

$$t_i(\theta_i) \equiv E_{\theta_{-i}}[t_i(\theta)].$$

Note that there is enormous flexibility in choosing actual transfers $t_i(\theta)$ to attain a specific $t_i(\theta_i)$ for implementability. Thus, in a truth-telling Bayesian-Nash equilibrium,

$$U_i(\hat{\theta}_i|\theta_i) \equiv E_{\theta_{-1}}[U_i(\theta_{-i},\hat{\theta}_i|\theta_i)] = \phi_i(\hat{\theta}_i)\theta_i - t_i(\hat{\theta}_i).$$

Let $U_i(\theta_i) \equiv U_i(\hat{\theta}_i|\theta_i)$. Following Theorem 3.2.2, we have

Theorem 26 An auction mechanism with $\phi_i(\theta_i)$ continuous and absolutely continuous first derivative is incentive compatible iff

$$\frac{dU_i(\theta_i)}{d\theta_i} = \phi_i(\theta_i),$$

and $\phi_i(\theta_i)$ is nondecreasing in θ_i.

The proof proceeds as in in Theorem 3.2.2. We are now in a position to examine the expected value of an incentive compatible auction.

The principal's objective function is to maximize expected payments taking into account the principal's own value of the object, which we take to be θ_0. Specifically,

$$\max_{\phi\in\Delta(I),t} E_\theta\left[\left(1-\sum_{i=1}^{I}\phi_i(\theta)\right)\theta_0 + \sum_{i=1}^{I}\phi_i(\theta)\theta_i - \sum_{i=1}^{I}U(\theta)\right],$$

subject to IC and IR. $\Delta(I)$ is the $I-1$ dimensional simplex.

As before, we can substitute out the indirect utility functions by integrating by parts. We are thus left the following objective function:

$$E_\theta\left[\left(1-\sum_{i=1}^{I}\phi_i(\theta)\right)\theta_0 + \sum_{i=1}^{I}\phi_i(\theta)\theta_i - \sum_{i=1}^{I}\left(\phi_i(\theta)\frac{1-P_i(\theta_i)}{p_i(\theta_i)} + U_i(0)\right)\right].$$

Rearranging the expression, we obtain

$$E_\theta\left[\theta_0 + \sum_{i=1}^{I}\phi_i(\theta)\left(\theta_i - \frac{1-P_i(\theta_i)}{p_i(\theta_i)} - \theta_0\right) - \sum_{i=1}^{I}U_i(0)\right]. \tag{9}$$

This last expression states the expected value of the auction independently of the transfer function. The expected value of the auction is completely determined by ϕ and $U(0) \equiv (U_1(0),\ldots,U_I(0))$. Any two auctions with the same functions have the same expected revenue.

Theorem 27 Revenue Equivalence. The seller's expected utility from an implementable auction is completely determined by the probability functions, ϕ, and the numbers, $U_i(0)$.

The proof follows from inspection of the expected revenue function, (9). The result is quite powerful. Consider the case of symmetric distributions of types, $p_i \equiv p$. The result implies that in the class of auctions which award the good to the highest value bidder and leave no rents to the lowest possible bidder (i.e., $U_i(0) = 0$), the expected revenue to the seller is the same. With the appropriately chosen reservation prices, the first-price auction, the second-price auction, the Dutch auction and the English auction all belong to this class! This extends Vickrey's [1961] famous equivalence result. Note that these auctions are not always optimal, however.

Back to optimality. Because the problem requires that $\phi \in \Delta(I-1)$, it is likely that we have a corner solution which prevents us from using first-order calculus techniques. As such, we do not redefine Φ and check A.2 and A.3. Instead, we will solve the problem directly. To this end, define

$$J_i(\theta_i) \equiv \theta_i - \frac{1 - P_i(\theta_i)}{p_i(\theta_i)}.$$

This is Myerson's virtual utility or virtual type for agent i with type θ_i. The principal thus wishes to maximize

$$E_\theta \quad \sum_{i=1}^{I} \phi_i(\theta) \left(J_i(\theta_i) - \theta_0\right) - U_i(0) \quad ,$$

subject to $\phi \in \Delta(I)$, monotonicity and $U_i(\theta_i) \geq 0$. We now state our result.

Theorem 28 Assume that each P_i satisfies the MHRC. Then the optimal auction has ϕ chosen such that

$$\phi_i(\theta) = \begin{cases} 1 & \text{if } J_i(\theta_i) > \max_{k \neq i} J_k(\theta_k) \text{ and } J_i(\theta_i) \geq \theta_0, \\ \in [0,1] & \text{if } J_i(\theta_i) = \max_{k \neq i} J_k(\theta_k) \text{ and } J_i(\theta_i) \geq \theta_0, \\ 0 & \text{otherwise.} \end{cases}$$

The lowest types receive no rents, $U_i(0) = 0$, and transfers satisfy the differential equation in Theorem 26.

Proof: Note first that the choice of ϕ in the theorem satisfies $\phi \in \Delta(I-1)$ and maximizes the value of (9). The choice of $U_i(0) = 0$ satisfies the participation constraints of the agents while maximizing profits. Lastly, the transfers are chosen so as to satisfy the differential equation in Theorem 26. Providing that $\phi_i(\theta_i)$ is nondecreasing, this implies incentive compatibility. To see that this monotonicity holds, note that $\phi_i(\theta)$ (weakly) increases as $J_i(\theta_i)$ increases, holding all other θ_{-i} fixed. The assumption of MHRC implies that $J_i(\theta_i)$ is increasing in θ_i, which implies the necessary monotonicity. □

The result is that the optimal auction awards the good to the agent with the highest virtual type, providing that the type exceeds the seller's opportunity cost, θ_0.

Remarks:

1. There are two distortions which the principal introduces, underconsumption and misallocation. First, sometimes the good will not be consumed even though an agent values it more than the principal:

   $$\max_i J_i(\theta_i) < \theta_0 < \max_i \theta_i.$$

   Second, sometimes the wrong agent will consume the good:

   $$\arg\max_i J_i(\theta_i) \neq \arg\max_i \theta_i.$$

2. The expected revenue of the optimal auction normally exceeds that of the standard English, Dutch, first-price, and second-price auctions. One reason is that if the distributions are discrete, the principal can elicit truth-telling more cheaply.

   Second, unless type distributions are symmetric, the J_i functions are asymmetric, which in turn implies that the highest value agent should not always get the item. In the four standard auctions, the highest valuation agent typically gets the good. By handicapping agents with more favorable distributions, however, you can encourage them to bid more aggressively. For example, let $\theta_0 = 0$, $I = 2$, θ_1 be uniformly distributed on $[0, 2]$ and θ_2 be uniformly distributed on $[2, 4]$. We have $J_1(\theta_1) \equiv 2(\theta_1 - 1)$ and $J_2(\theta_2) \equiv 2\theta_2 - 4 = J_1(\theta_2) - 2$. Agent 2 is handicapped by 2 relative to agent 1 in this mechanism (i.e., agent 1 must have a value in excess of agent 2 by at least 2 in order to get the good). In contrast, under a first-price auction, agent 2 always gets the good, submitting a bid of only 2.

3. With correlation, the principal can do even better. See Myerson [1981], for an example, and Crémer and McLean [1985,1988] for more details. There is also a literature beginning with Milgrom and Weber [1982] on common value (more precisely, affiliated value) auctions.

4. With risk aversion, the revenue equivalence theorem fails to apply. Here, first-price generally outperforms second-price, for example. The idea is that if you are risk averse, you will bid more aggressively in a first-price auction because bidding close to your valuation reduces the risk in your final rent. See Maskin-Riley [1984], Matthews [1983] and Milgrom and Weber [1982] for more on this subject.

5. Maskin and Riley [1990] consider multi-unit auctions where agents have multi-unit demands. The result is a combination of the above result on virtual valuations allocation and the result from price-discrimination regarding the level of consumption for those who consume in the auction.

6. Although the revenue equivalence theorem tells us that the four standard auctions are equivalent in revenue, it is in the context of Bayesian implementation. Note, however, the second-price sealed bid auction has a unique dominant-strategy equilibrium. Thus, revenue may not be the only relevant dimension over which we should value a mechanism.

Bilateral (Multilateral) Trading This section is based upon Myerson-Satterthwaite [1983]. Here, we reproduce their characterization of implementable trading mechanisms between a buyer and a seller with privately known valuations of trade. We then derive the nature of the optimal mechanism. We put off until later the discussion of efficiency properties of mechanisms with balanced budgets.

Basic Model of a Double Auction:

The basic bilateral trade model of Myerson and Satterthwaite has two agents: a seller and a buyer. There is no principal. Instead think of the optimization problem as the design of a mechanism by the buyer and seller before they know their types, but under the condition that after learning their types, either party may walk away from the agreement.

Additionally, it is assumed that money can only be transferred from one party to the other. There is no outside party that can break the budget. The seller and the buyer's valuations for the single unit of good are $c \in [\underline{c}, \overline{c}]$ and $v \in [\underline{v}, \overline{v}]$ respectively; the distributions are $P_1(c)$ for the seller and $P_2(v)$ for the buyer. [We'll use v and c rather than the θ_i as they're more descriptive.]

An allocation is given by $y = (\phi, t)$ where $\phi \in [0, 1]$ is the probability of trade and t is a transfer from the buyer to the seller. Thus, the indirect utilities are

$$U_1(\hat{c}, v|c) \equiv t(\hat{c}, v) - \phi(\hat{c}, v)c,$$

and

$$U_2(c, \hat{v}|v) \equiv \phi(c, \hat{v})v - t(c, \hat{v}).$$

Using the appropriate expectations, in a truth-telling equilibrium we have

$$U_1(\hat{c}|c) \equiv t(\hat{c}) - \phi(\hat{c})c,$$

and

$$U_2(\hat{v}|v) \equiv \phi(\hat{v})v - t(\hat{v}).$$

Characterization of Implementable Contracts: M&S provide the following useful characterization theorem.

Theorem 29 For any probability function $\phi(c, v)$, there exists a transfer function t such that $y = (\phi, t)$ is IC and IR iff

$$E_{v,c} \quad \phi(c,v) \left\{ \left(v - \frac{1 - P_2(v)}{p_2(v)} \quad - \quad c + \frac{P_1(c)}{p_1(c)} \right) \right\} \Bigg] \geq 0, \tag{10}$$

$\phi(v)$ is nondecreasing, and $\phi(c)$ is nonincreasing.

Sketch of Proof: The proof of this claim is straightforward. Necessity follows from our standard arguments. Note first that substituting out the transfer function from the two indirect utility functions (which was can do sense the transfers must be equal under budget balance) and taking expectations implies

$$E_{c,v}[U_1(c) + U_2(v)] = E_{c,v}[\phi(c,v)(v - c)].$$

Using the standard arguments presented above, one can show that this is also equivalent to

$$E_{c,v} \quad U_1(\overline{c}) + \phi(c,v) \quad \frac{P_2(c)}{p_2(c)} \quad + U_2(\underline{v}) + \phi(c,v) \quad \frac{1 + P_1(v)}{p_1(v)} \quad .$$

Rearranging the expression and imposing individual rationality implies (10). Monotonicity is proved using the standard arguments. Sufficiency is a bit trickier. It involves finding the solution to the partial differential equations (first-order conditions) for incentive compatibility. This solution, together with monotonicity, is sufficient for truth-telling. See M&S for the details.

We now are prepared to find the optimal bilateral trading mechanism.

Optimal Bilateral Trading Mechanisms:

The "principal" wants to maximize the expected gains from trade,

$$E_{c,v}[\phi(c,v)(v-c)],$$

subject to monotonicity and (10). We will ignore monotonicity and check our solution to see that it is satisfied. Let μ be the constraint (10). Bringing the constraint into the objective function and simplifying, we have

$$\max_{c,v} E_{c,v}\left[\phi(c,v)\left\{(v-c) - \frac{\mu}{1+\mu}\left(\frac{1-P_2(v)}{p_2(v)} - \frac{P_1(c)}{p_1(c)}\right)\right\}\right].$$

Notice that trade occurs in this relaxed program iff

$$v - \frac{\mu}{1+\mu}\left(\frac{1-P_2(v)}{p_2(v)}\right) \geq c + \frac{\mu}{1+\mu}\left(\frac{P_1(c)}{p_1(c)}\right),$$

where $\mu \geq 0$. If we assume that the monotone hazard-rate condition is satisfied for both type distributions, then this ϕ is appropriately monotonic and we have a solution to the full program. Note that if $\mu > 0$, there will generally be inefficiencies in trade. This will be discussed below.

Remarks:

1. Importantly, M&S show that when $\overline{c} > \underline{v}$ and $\overline{v} > \underline{c}$ (i.e., efficient trading is state dependent), $\mu > 0$, so the full-information efficient level of trading is impossible! The proof is to show that substituting the efficient ϕ into the constraint (10) violates the inequality.

2. Chatterjee and Samuelson [1983] show that in a simple game in which each agent (buyer and seller) simultaneously submits an offer (i.e., a price at which to buy or to sell) and trade takes place at the average price iff the buyer's offer exceeds the seller's offer, if types are uniformly distributed then a Nash equilibrium in linear bidding strategies exists and achieves the upper bound for bilateral trading established by Myerson and Satterthwaite.

3. A generalization of this result to multi-lateral bargaining contexts is found in Cramton, Gibbons, and Klemperer [1987]. They look at the dissolution of a partnership, in which unlike the buyer-seller example where the seller owned the entire good, each member of the partnership may have some property claims on the partnership. The question is whether the partnership can be dissolved efficiently (i.e., the highest value partner buying out all other partners). They show that if the ownership is sufficiently distributed, that it is indeed possible.

Feasible Allocations and Efficiency There is an old but important literature concerning the implementation of optimal public choice rules, such as when to build a bridge. Bridges cost money, so it is efficient to build one only if the sum of the individual agents' values exceed the costs. The question is how to get agents to truthfully state their valuations (e.g., no one exaggerates their value to change the probability of building the bridge to their own benefit); i.e., how do you avoid the classical "free-rider" problem.

Three important results exist. First, if one ignores budget balance and individual rationality constraints, it is possible to implement the optimal public choice rule in dominant strategies. This is the contribution of Clarke [1971] and Groves [1973]. Second, if one requires budget balance, one can still implement the optimal rule if one uses Bayesian-Nash implementability. This is the result of d'Aspremont and Gerard-Varet [1979]. Finally, if one wants budget balance and individual rationality, efficient allocation is not generally possible even under the Bayesian-Nash concept, as shown by Myerson and Satterthwaite's result [1983]. We consider the first two results now.

The basic model is that there are I agents, each with utility function $u_i(x,\theta_i)+t_i$, where x is the decision variable. Some of these agents actually build the bridge, so their utility may depend negatively on the value of x. Let $x^*(\theta)$ be the unique solution to $\max_x \sum_{i=1}^I u_i(x,\theta_i)$.

To be clear, the various sorts of constraints are:

$$
\begin{aligned}
&\text{(Ex post BB)} && \sum_{i=1}^{I} t_i(\theta) \le 0,\ \forall\theta \\
&\text{(Ex ante BB)} && \sum_{i=1}^{I} E_\theta[t_i(\theta)] \le 0,\ \forall\theta \\
&\text{(BN-IC)} && E_{\theta_{-i}}[U_i(\theta_i,\theta_{-i}|\theta_i)] \ge E_{\theta_{-i}}[U_i(\hat\theta_i,\theta_{-i}|\theta_i)],\ \forall(\theta_i,\hat\theta_i) \\
&\text{(DS-IC)} && U_i(\theta_i,\theta_{-i}|\theta_i)] \ge U_i(\hat\theta_i,\theta_{-i}|\theta_i),\ \forall(\theta_i,\hat\theta_i,\theta_{-i}), \\
&\text{(Ex post IR)} && U_i(\theta) \ge 0\ \forall\theta, \\
&\text{(Interim IR)} && E_{\theta_{-i}}[U_i(\theta_i,\theta_{-i})] \ge 0,\ \forall\theta_i, \\
&\text{(Ex ante IR)} && E_\theta[U_i(\theta)] \ge 0,\ \forall\theta_i.
\end{aligned}
$$

The Groves Mechanism:

We want to implement x^* using transfers that satisfy DS-IC. The trick to implementing x^* in dominant strategies is to choose a transfer function for agent i that makes agent i's payoff equal to the social surplus.

$$t_i(\hat\theta) \equiv \sum_{j\ne i} u_j(x^*(\hat\theta_i,\hat\theta_{-i}),\hat\theta_j) + \tau_i(\hat\theta_{-i}),$$

where τ_i is an arbitrary function of θ_{-i}. To see that this mechanism $y=(x^*,t)$ is dominant-strategy incentive compatible, note that for any θ_{-i}, agent i's utility is

$$U_i(\hat\theta_i,\theta_{-i}|\theta_i) = \sum_{j=1}^{I} u_j(x^*(\hat\theta_i,\theta_{-i}),\theta_j) + \tau_i(\hat\theta_{-i}).$$

τ_i can be ignored because it is independent of agent i's report. Thus, agent i chooses $\hat{\theta}_i$ to maximize

$$U_i(\hat{\theta}_i, \theta_{-i}|\theta_i) = \sum_{j=1}^{I} u_j(x^*(\hat{\theta}_i, \theta_{-i}), \theta_j).$$

By definition of $x^*(\theta)$, the choice of $\hat{\theta}_i = \theta_i$ is optimal *for any* θ_{-i}.

Green and Laffont [1977] have shown that *any mechanism with truth-telling as a dominant strategy has the form of a Grove's mechanism.* Also, they show that in general, ex post BB is violated by a Grove's mechanism. Given the desirability of BB, we turn to a less powerful implementation concept.

The AGV Mechanism:

d'Aspremont and Gerard-Varet (AGV) [1979] show that x^* can be implemented with transfers satisfying BB if one only requires BN-IC (rather than DS-IC) to hold.

Consider the transfer

$$t_i(\hat{\theta}) \equiv E_{\theta_{-i}} \left[\sum_{j \neq i} u_j(x^*(\theta_{-i}, \hat{\theta}_i), \theta_j) \right] + \tau_i(\hat{\theta}_{-i}).$$

τ_i will be chosen to ensure BB is satisfied. Note that the agent's expected payoff given that $\hat{\theta}$ is announced is (ignoring τ_i)

$$E_{\theta_{-i}} \left[u_i(x^*(\hat{\theta}_i, \theta_{-i}), \theta_i) + \sum_{j \neq i} u_j(x^*(\theta_{-i}, \hat{\theta}_i), \theta_j) \right].$$

Providing that all the other players announce truthfully, $\hat{\theta}_{-i} = \theta_{-i}$, player i's optimal strategy is truth-telling. Hence, BN-IC is satisfied by the transfers.

Now we construct τ_i so as to achieve BB.

$$\tau_i(\hat{\theta}_{-i}) \equiv -\frac{1}{I-1} \sum_{j \neq i} E_{\theta_{-j}} \left[\sum_{k \neq j} u_k(x^*(\hat{\theta}_j, \theta_{-j}), \theta_k) \right].$$

Intuitively, the τ_i are constructed so as to have i pay off portions of the other players' subsidies (which are independent of i's report).

Remarks:

1. If the budget constraint were instead $\sum_{i=1}^{N} t_i \leq C_0$, where C_0 is the cost of provision, the same argument of budget balance for the AGV mechanism can be made.

2. Note that the AGV mechanism does not guarantee that ex post or interim individual rationality will be met. Ex ante IR can be satisfied with appropriate side payments.

3. The AGV mechanism is very useful. Suppose that two individuals can write a contract *before* learning their types. We call this the *ex ante* stage in an incomplete information game. Then the players can use an AGV mechanism to get the first best tomorrow when they learn their types, and they can transfer money among themselves today to satisfy any division of the expected surplus that they like. In particular, they can transfer money ex ante to take into account that their IR constraints may be violated ex post. [This requires that the parties can commit not to walk away ex post if they lose money.] *Therefore, ex ante contracting with risk neutrality implies no inefficiencies ex post.* Of course, most contracting situations we have considered so far involve the contract being offered at the *interim stage* where each party knows their own information, but not the information of others. At this stage we frequently have distortions emerging because of the individual rationality constraints. At the *ex post stage*, all information is known.

4. As said above, in Myerson and Satterthwaite's [1983] model of bilateral exchange, mechanisms that satisfies ex post BB, BN-IC, and interim IR are inefficient when efficient trade depends upon the state of nature.

5. Note that Myerson and Satterthwaite's [1983] inefficiency result continues to hold if we require only interim BB rather than ex post BB. Interestingly, Williams [1995] has demonstrated that if the constraints on the bilateral trading problem are all ex ante or interim in nature, if utilities are linear in money, and if the first-best outcome is implementable in a BNE mechanism, then it is also implementable in a DSE mechanism (ala' Clark-Groves). In other words, the space of efficient BNE mechanisms is spanned by Clark-Groves mechanisms. Thus, in the interim BB version of Myerson-Satterthwaite, showing that the first-best is not implementable with a Clark-Groves mechanism is sufficient for demonstrating that the first-best is also not implementable with any BNE mechanism.

6. Some authors have extended M&S by considering many buyers and sellers in a market mechanism design setting. With double auctions, as the number of agents increases, trade becomes full-information efficient.

7. With many agents in a public goods context, the limit results are negative. As Mailath and Postlewaite [1990] have shown, as the number of agent's becomes large, an agent's information is rarely pivotal in a public goods decision, and so inducing the agent to tell the truth require subsidies that violate budget balance.

3.2.7 General Remarks on the Static Mechanism Design Literature

1. The timing discussed above has generally been that the principal offers a contract to the agent at the interim stage (after the agent knows his type). We have seen that an AGV mechanism can generally obtain the first best if unrestricted contracting can occur at the ex ante stage an the agents are risk neutral. Sappington [1983] considers an interesting *hidden information* model which demonstrates that if contracting occurs at the ex ante stage among risk neutral agents *but where the agents must be guaranteed some utility level in every state* (ex post IR), the contracts are similar to the standard

interim screening contracts. Thus, if agents can always threaten to walk away from the contract after learning their type, it is as if you are in the interim contracting game.

2. The above models generally assume an IR constraint which is independent of type. This is not always plausible as high type agents may have better outside options. When one introduces type-contingent IR constraints, it is no longer clear where the IR constraint binds. The models become messy but sometimes yield considerable new economic insight. There are a series of very nice papers by Lewis and Sappington [1989a,1989b] which investigate the use of countervailing incentives which endogenously use type-dependent outside options to the principal's advantage. A nice extension and unification of their approach is given by Maggi and Rodriguez [1995], which indicates the relationship between countervailing incentives and inflexible rules and how Lewis and Sappington's results depend importantly upon whether the agent's utility is quasi-concave or quasi-convex in the information parameter. The most general paper on the techniques involved in designing screening contracts when agent's utilities depend upon their type is given by Jullien [1995]. Applications using countervailing incentives such as Laffont-Tirole [1990] and Stole [1995] consider the effects of outside options on the optimal contracts and find whole intervals of types in which the IR constraints bind.

3. Another interesting extension of the basic paradigm is the construction of general equilibrium models that endogenously determine outside options. This, combined with type-contingent outside options, has been studied by Spulber [1989] and Stole [1995] in price discrimination contexts.

3.3 Dynamic Principal-Agent Screening Contracts

We now turn to the case of dynamic relationships. We restrict our attention to two period models where the agent's relevant private information does not change over time in order to keep things simple. There are three environments to consider, each varying in terms of the commitment powers of the principal. First, there is full-commitment, where the principal can credibly commit to an incentive scheme for the duration of the relationship. Second, there is the other extreme, where no commitment exists and contracts are effectively one-period contracts: any two-period contract can be torn up by *either party* at the start of the second-period. This implies in particular that the principal will always offer a second period contract that optimally utilizes any information revealed in the first period. Third, midway between these regimes, is the case of commitment with renegotiation: Contracts cannot be torn up unless *both parties* agree to it, thus renegotiations must be Pareto improving. We consider each of these environments in turn.

This section is largely based upon Laffont and Tirole [1988,1990], which are nicely presented in Chapters 9 and 10 of their 1993 book. The main difference is in the economic setting; here we study nonlinear pricing contracts (ala' Mussa-Rosen [1978]) rather than the procurement environment. Also of interest for the third regime of renegotiation with commitment discussed in subsection 3.3.4 is Dewatripont [1989]. Lastly, Hart and Tirole [1988] offer another study of dynamic screening contracts across these three regimes in an intertemporal price discrimination framework, discussed below in section 3.3.5.

3.3.1 The Basic Model

We will utilize a simple model of price discrimination (ala' Mussa-Rosen [1978]) throughout the analysis but our results do not depend on this model directly. [Laffont-Tirole [1993; Chapters 1, 9 and 10] perform a similar analysis in the context of their regulatory framework.] Suppose that the firm's unit cost of producing a product with quality q is $C(q) \equiv \frac{1}{2}q^2$. A customer of type θ values a good of quality q by $u(q,\theta) \equiv \theta q$. Utilities of both actors are linear and transferable in money: $V \equiv t - C(q)$, $U \equiv \theta q - t$, and $S(q,\theta) \equiv \theta q - C(q)$. We will consider two different informational settings. First, the continuous case, where θ is distributed according to $P(\theta)$ on $[\underline{\theta}, \overline{\theta}]$; second, the two-type case, where $\overline{\theta}$ occurs with probability p and $\underline{\theta}$ occurs with probability $1-p$. Under either setting, the first-best full-information solution is to set $q(\theta) = \theta$. Under a one-period private information setting, our results for the choice of quality under each setting are, for the continuous case,

$$q(\theta) = \theta - \frac{1 - P(\theta)}{p(\theta)}, \; \forall\, \theta,$$

and for the two-type case,

$$\overline{q} \equiv q(\overline{\theta}) = \overline{\theta},$$

and

$$\underline{q} \equiv q(\underline{\theta}) = \underline{\theta} - \frac{p}{1-p}\Delta\theta,$$

where $\Delta\theta \equiv \overline{\theta} - \underline{\theta}$. We assume throughout that $\underline{\theta} > p\overline{\theta}$, so that the firm wishes to serve the agent (i.e., $q(\underline{\theta}) \geq 0$).

3.3.2 The Full-Commitment Benchmark

Suppose that the relationship lasts for two contracting periods, where the common discount factor for the second period is $\delta > 0$. We allow either $\delta < 1$ or $\delta > 1$ to generally reflect the relative importance of the payoffs in the two periods. We have the following immediate result.

Theorem 30 Under full-commitment, the optimal long-term contract commits the principal to offering the optimal static contract in each period.

Proof: The proof is simple. Because $u_{\theta qq} = 0$, there is no value to introducing randomization in the static contract. In particular, there is no value to offering one allocation with probability $\frac{1}{1+\delta}$ and the other with probability $\frac{\delta}{1+\delta}$. But this also implies that the principal's contract should not vary over time. □

3.3.3 The No-Commitment Case

Now consider the other extreme of no commitment. We show two results. First, in the continuum case for any implementable first-period contract pooling must occur almost everywhere. Second, in the two-type case, although separation may be feasible, separation is generally not optimal and some pooling will be induced in the first period. Of particular interest is that in the second case, both upward and downward incentive compatibility constraints may bind because the low-type agent can always "take-the-money-and-run."

Continuous Case:

With a continuum of types, it is impossible to get separation almost everywhere. Let $U_2(\hat{\theta}|\theta)$ be the continuation equilibrium payoff that type θ obtains in period 2 given that the principal believes the agent's type is actually $\hat{\theta}$. Note that if there is full separation for some type θ, then $U_2(\theta|\theta) = 0$. Additionally, if there is separation in equilibrium between θ and $\hat{\theta}$, then out of equilibrium, $U_2(\hat{\theta}|\theta) = \max\{(\theta - \hat{\theta})q(\hat{\theta}), 0\}$.

The basic idea goes like this. Suppose that type θ is separated from type $\hat{\theta} = \theta - d\theta$ in the period 1 contract. By lying downward in the first period, the high type suffers only a second-order loss; but being thought to be a lower type in the second period raises the high type's payoff by a first-order amount: $\delta U_2(\theta - d\theta|\theta)$. Specifically, we have the following theorem.

Theorem 31 For *any* first-period contract, there exists no non-degenerate subinterval of $[\underline{\theta}, \overline{\theta}]$ in which full separation occurs.

Proof: Suppose not and there is full sorting over $(\theta_0, \theta_1) \subset [\underline{\theta}, \overline{\theta}]$.

Step 1. (q, t) increases in θ implying that the functions are almost everywhere differentiable. Take $\theta > \hat{\theta}$, both in the subinterval. By incentive compatibility,

$$\theta q(\theta) - t(\theta) \geq \theta q(\hat{\theta}) - t(\hat{\theta}) + \delta U_2(\hat{\theta}|\theta),$$

$$\hat{\theta} q(\hat{\theta}) - t(\hat{\theta}) \geq \hat{\theta} q(\theta) - t(\theta).$$

The first equation represents that the high type agent will receive a positive rent from deceiving the principal in the first period. This rent term is not in the second equation because along the equilibrium path (truthtelling), no agent makes rents in the second period, and so the lower type will prefer to quit the relationship rather than consume the second-period bundle for the high-type which would produce negative utility for the low type. [This is the "take-the-money-and-run" strategy.] Given the positive rent from lying in the second period, adding these equations together imply $(\theta - \hat{\theta})(q(\theta) - q(\hat{\theta})) > 0$, which implies that $q(\theta)$ is strictly increasing over the subinterval (θ_0, θ_1). This result, combined with either inequality above, implies that $t(\theta)$ is also strictly increasing.

Step 2. Consider a point of differentiability, θ, in the subinterval. Incentive compatibility between θ and $\theta + d\theta$ implies

$$\theta q(\theta) - t(\theta) \geq \theta q(\theta + d\theta) - t(\theta + d\theta),$$

or

$$t(\theta + d\theta) - t(\theta) \geq \theta [q(\theta + d\theta) - q(\theta)].$$

Dividing by $d\theta$ and taking the limit as $d\theta \to 0$ yields

$$\frac{dt(\theta)}{d\theta} \geq \theta \frac{dq(\theta)}{d\theta}.$$

Now consider incentive compatibility between θ and $\theta - d\theta$ for the type θ agent. An agent of type θ who is mistaken for an agent of type $\theta - d\theta$ will receive second period rents of $U_2(\theta - d\theta|\theta) = q(\theta - d\theta)d\theta > 0$. As a consequence,

$$\theta q(\theta) - t(\theta) \geq \theta q(\theta - d\theta) - t(\theta - d\theta) + \delta U_2(\theta - d\theta|\theta),$$

or

$$t(\theta) - t(\theta - d\theta) \leq \theta[q(\theta) - q(\theta - d\theta)] - \delta q(\theta - d\theta)d\theta,$$

or taking the limit as $d\theta \to 0$,

$$\frac{dt(\theta)}{d\theta} \leq \theta \frac{dq(\theta)}{d\theta} - \delta q(\theta).$$

Combining the two inequalities yields

$$\theta \frac{dq(\theta)}{d\theta} \leq \frac{dt(\theta)}{d\theta} \leq \theta \frac{dq(\theta)}{d\theta} - \delta q(\theta),$$

which is a contradiction. □

In their paper, Laffont-Tirole [1988] also characterize the nature of equilibria (in particular, they provide necessary and sufficient conditions for partition equilibria for quadratic utility functions in their regulatory context). Instead of studying this, we turn to the simpler two-type case.

Two-type Case:

Before we begin, some remarks are in order.

Remarks:

1. Note a very important distinction between the continuous case and the two-type case. For the *continuous type case*, full separation over any subinterval *is not implementable*; in the *two-type case* we will see that separation may be implementable but typically it *is not optimal.*

2. Because we shall look at the principal's optimal choice of contracts, we need to make clear the notion of continuation equilibria. At the end of the first period, several equilibria may exist in the continuation of the game. We will generally consider the best outcome from the principal's point of view, but if one is worried about uniqueness, one needs to be more careful here.

3. We will restrict our attention to menus with only two contracts. We do not know if this is without loss of generality.

4. We will also restrict our attention to parameters values such that the principal always prefers to serve both types in each period.

5. Let ν be the principal's probability assessment that the agent is of type $\overline{\theta}$. Then the conditionally optimal contract has $\overline{q} = \overline{\theta}$ and $\underline{q} = \underline{\theta} - \frac{\nu}{1-\nu}\Delta\theta$. Thus, the low-type's allocation decreases with the principal's assessment.

We will proceed by restricting our attention to two-part menus. There will be three cases of interest. We will analyze the optimal contracts for these cases.

Consider the following contract offer: $(\underline{q}_1, \underline{t}_1)$ and $(\overline{q}_1, \overline{t}_1)$. We denote periods by the subscript on the contracting variables. Without loss of generality, assume that $\underline{\theta}$ chooses $(\underline{q}_1, \underline{t}_1)$ with positive probability and $\overline{\theta}$ chooses $(\overline{q}_1, \overline{t}_1)$ with positive probability. Let $U_2(\nu|\overline{\theta})$ be the rent the high type receives in the continuation game where the principal's belief that the agent is of type $\overline{\theta}$ is ν. We will let $\overline{\nu}$ represent the belief of the principal after observing the contract choice $(\overline{q}_1, \overline{t}_1)$ and $\underline{\nu}$ represent the belief of the principal after observing the contract choice $(\underline{q}_1, \underline{t}_1)$. Then the principal will design an allocation which maximizes profit subject to four constraints.

$$
\begin{array}{lrcl}
(\overline{IC}) & \overline{\theta}\overline{q}_1 - \overline{t}_1 + \delta U_2(\overline{\nu}|\overline{\theta}) & \geq & \overline{\theta}\underline{q}_1 - \underline{t}_1 + \delta U_2(\underline{\nu}|\overline{\theta}), \\
(\underline{IC}) & \underline{\theta}\underline{q}_1 - \underline{t}_1 + \delta U_2(\underline{\nu}|\underline{\theta}) & \geq & \underline{\theta}\overline{q}_1 - \overline{t}_1 + \delta U_2(\overline{\nu}|\underline{\theta}), \\
(\overline{IR}) & \overline{\theta}\overline{q}_1 - \overline{t}_1 + \delta U_2(\overline{\nu}|\overline{\theta}) & \geq & 0, \\
(\underline{IR}) & \underline{\theta}\underline{q}_1 - \underline{t}_1 & \geq & 0.
\end{array}
$$

As is usual, $\overline{IR}$ is implied by the $\overline{IC}$ and $\underline{IR}$. Additionally, $\underline{IR}$ must be binding (providing the principal gets to choose the continuation equilibrium). To see this, note that if it were not binding, both $\underline{t}$ and $\overline{t}$ could be raised by equal amounts without violating the IC constraints and profits could be increased. Thus, the principal can substitute out $\underline{t}$ from the objective function using $\underline{IR}$ (thereby imposing this constraint with an equality). Now the principal's problem is to maximize profits subject to the two IC constraints.

There are three cases to consider. Contracts in which only the high-type's IC constraint is binding (Type I); contracts in which only the low-type's IC constraint is binding (Type II); and contracts in which both IC constraints bind (Type III). In turns out that Type II contracts are never optimal for the principal, so we will ignore them. [See Laffont and Tirole, 1988, for the argument.] Type I contracts are the simplest to study; Type III contracts are more complicated because the take-the-money-and-run strategy of the low type causes the low-type's IC constraint to bind upward.

Type I Contracts: Let the $\overline{\theta}$ type customer choose the high-type contract, $(\overline{q}_1, \overline{t}_1)$ with probability $1-\alpha$ and $(\underline{q}_1, \underline{t}_1)$ with probability α. In a type I equilibrium, whenever the $(\overline{q}_1, \overline{t}_1)$ contract is chosen, the principal's beliefs are degenerate: $\overline{\nu} = 1$. Following Bayes' rule, when $(\underline{q}_1, \underline{t}_1)$ is chosen, $\underline{\nu} = \frac{\alpha p}{\alpha p + (1-p)} < p$. Note that $\frac{d\nu}{d\alpha} > 0$. Define the single-period expected asymmetric information profit level of the principal with belief ν who offers the optimal static contract as

$$\Pi(\nu) \equiv \max_{\overline{q},\underline{q}} \ \nu(\overline{\theta}\overline{q} - C(\overline{q}) - \Delta\theta\underline{q}) + (1-\nu)(\underline{\theta}\underline{q} - C(\underline{q})).$$

Consider the second period. Given any belief ν, the principal will choose the conditionally optimal contract. This implies a choice of $\overline{q}_2 = \overline{\theta}$, $\underline{q}_2(\alpha) = \underline{\theta} - \alpha\frac{p}{1-p}\Delta\theta$, and profit

$\Pi(\underline{\nu}(\alpha))$, where we have directly acknowledged the dependence of $\underline{q}_2$ and second period profits on α.[5] Thus, we are left with calculating the first-period choices of $(\overline{q}_1, \underline{q}_1, \alpha)$ by the principal. Specifically, the principal solves

$$\max_{(\overline{q}_1,\overline{t}_1),(\underline{q}_1,\underline{t}_1),\alpha} p\left\{(1-\alpha)[\overline{t}_1 - C(\overline{q}_1) + \delta\Pi(1)] + \alpha[\underline{t}_1 - C(\underline{q}_1) + \delta\Pi(\underline{\nu}(\alpha))]\right\}$$
$$+ (1-p)[\underline{t}_1 - C(\underline{q}_1) + \delta\Pi(\underline{\nu}(\alpha))],$$

subject to

$$\overline{\theta}\overline{q}_1 - \overline{t}_1 = \overline{\theta}\underline{q}_1 - \underline{t}_1 + \delta\Delta\theta\underline{q}_2(\alpha),$$

and

$$\underline{\theta}\underline{q}_1 - \underline{t}_1 = 0.$$

The first constraint is the binding IC constraint for the high type where $U_2(\underline{\nu}(\alpha)|\overline{\theta}) = \Delta\theta\underline{q}_2(\alpha)$ given our previous discussion of price discrimination. The second constraint is the binding IR constraint for the low type. Note that by increasing α, the principal directly decreases profit by learning less information, but simultaneously credibly lowers $\underline{q}_2$ in the second period which weakens the high-type's IC constraint in the first period. Thus, there is a tradeoff between separation (which is good per se because this improves efficiency) and the rents which must be given to the high type to obtain the separation. Substituting the constraints into the principal's objective function and simplifying yields

$$\max_{\overline{q}_1,\underline{q}_1,\alpha} p(1-\alpha)[\overline{\theta}\overline{q}_1 - C(\overline{q}_1) - \Delta\theta\underline{q}_1 - \delta\Delta\theta\underline{q}_2(\alpha)]$$
$$+ (1-p+\alpha p)[\underline{\theta}\underline{q}_1 - C(\underline{q}_1)] + p(1-\alpha)\delta\Pi(1) + (1-p+\alpha p)\delta\Pi(\underline{\nu}(\alpha)).$$

The first-order conditions for output are $\overline{q}_1 = \overline{\theta}$ and

$$\underline{q}_1 = \underline{\theta} - \frac{p-\alpha p}{(1-p)+\alpha p}\Delta\theta.$$

The latter condition indicates that $\underline{q}_1$ will be chosen at a level *above* that of the static contract if $\alpha > 0$. To see the choice of $\underline{q}_1$ in a different way, note that by rearranging the terms of the above objective function we have

$$\underline{q}_1 = \arg\max_q \quad \alpha p(\overline{\theta}q - C(q)) + (1-p)(\underline{\theta}q - C(q)) - p\Delta\theta q.$$

Here, $\underline{q}_1$ is chosen taking into account the efficiency costs of pooling and the rent which must be left to the high type.

Finally, one must maximize subject to α, which will take into account the tradeoff between surplus-increasing separation and reducing the rents which the high-type receives in order to separate in period one with a higher probability. Generally, we will find that

[5] In general, it is not enough to assume that $\Delta\theta$ is small enough that the principal always prefers to serve both types in the static equilibrium, because we might suspect that in a dynamic model low types may not be served in a later period. For Type I contracts, this is not in issue as $\underline{\nu} < p$, so low types are even more attractive in the second period than in the first when $(\underline{q}_1, \underline{t}_1)$ is chosen. When the other contract is chosen, no low types exist, and so this is irrelevant. Unfortunately, this is not the case with Type III contracts.

the pooling probability, α, is increasing in δ. Thus, as the second-period becomes more important, less separation occurs in the first period.

One can demonstrate that if δ is sufficiently large, the IC constraint for the low type will be binding, and so one must check the solution to the Type I program to verify that it is indeed a type I equilibrium (i.e., the IC constraint for the low type is slack at the optimum). For high δ, this will not be the case, and so we have a type III contract.

Type III Contracts:

Let the high type choose $(\underline{q}_1, \underline{t}_1)$ with probability α as before, but let the low type choose $(\overline{q}_1, \overline{t}_1)$ with probability β. Now, by Bayes' rule, we have

$$\overline{\nu}(\alpha, \beta) \equiv \frac{p(1-\alpha)}{p(1-\alpha) + (1-p)(1-\beta)},$$

and

$$\underline{\nu}(\alpha, \beta) \equiv \frac{p\alpha}{p\alpha + (1-p)\beta}.$$

As before, the second-period contract will be conditionally optimal. This implies that $\overline{q}_2 = \overline{\theta}$, $\underline{q}_2(\nu) = \underline{\theta} - \frac{\nu}{1-\nu}\Delta\theta$, and $U_2(\nu|\overline{\theta}) = \Delta\theta \underline{q}_2(\nu)$.

The principal's type III program is

$$\max \; [p(1-\alpha) + (1-p)\beta][\overline{t}_1 - C(\overline{q}_1) + \delta\Pi(\overline{\nu}(\alpha, \beta)] \\ + [p\alpha + (1-p)(1-\beta)][\underline{t}_1 - C(\underline{q}_1) + \delta\Pi(\underline{\nu}(\alpha, \beta)),]$$

subject to

$$\overline{\theta}\overline{q}_1 - \overline{t}_1 + \delta U_2(\overline{\nu}(\alpha, \beta)) = \overline{\theta}\underline{q}_1 - \underline{t}_1 + \delta U_2(\underline{\nu}(\alpha, \beta)),$$
$$\underline{\theta}\underline{q}_1 - \underline{t}_1 = \underline{\theta}\overline{q}_1 - \overline{t}_1,$$
$$\underline{\theta}\underline{q}_1 - \underline{t}_1 = 0.$$

The first and second constraints are the binding IC constraints; the third constraint is the binding IR constraint for the low type. Manipulating the three constraints implies that $\Delta\theta(\overline{q}_1 - \underline{q}_1) = \delta[U_2(\underline{\nu}(\alpha, \beta)) - U_2(\overline{\nu}(\alpha, \beta))]$. Substituting the three constraints into the objective function to eliminate t and simplifying, we have

$$\max_{\overline{q}_1, \underline{q}_1, \alpha, \beta} \; [p(1-\alpha) + (1-p)\beta][\underline{\theta}\overline{q}_1 - C(\overline{q}_1) + \delta\Pi(\overline{\nu}(\alpha, \beta)] \\ + [p\alpha + (1-p)(1-\beta)][\underline{\theta}\underline{q}_1 - C(\underline{q}_1) + \delta\Pi(\underline{\nu}(\alpha, \beta)),]$$

subject to

$$\overline{q}_1 - \underline{q}_1 = \delta[\underline{q}_2(\underline{\nu}(\alpha, \beta)) - \underline{q}_2(\overline{\nu}(\alpha, \beta))].$$

First-order conditions imply that $\overline{q}_1$ is not generally efficient because of the IC constraint for the low type. In fact, given Laffont and Tirole's simulations, it is generally possible that first-period allocation may be decreasing in type: $\overline{q}_1 < \underline{\theta} < \underline{q}_1$! [Note that a nondecreasing allocation is no longer a necessary condition for incentive compatibility because of the second period rents.]

Remarks for the Two-type Case (Type I and III):

1. For δ small, the low-type's IC constraint will not bind and so we will have a type I contract. For sufficiently high δ, the reverse is true.

2. As δ goes to ∞, $\underline{q}_1$ goes to $\overline{q}_1$ and there is complete pooling in the first period. The idea is that by pooling in the first period, the principal can commit not to learn anything and therefore impose the statically optimal separation contract in the second period.

3. Note that for any finite δ, complete pooling in period 1 is never optimal. That is, the above result is a limit result only. To see why, suppose that full pooling were undertaken. Then the second period output levels are statically optimal. By reducing pooling in the first period by a small degree, surplus is increased by a first-order amount while there is only a second-order effect on profits in the second period (since it was previously statically optimal).

4. Clearly, the firm always prefers commitment to non-commitment. In addition, for δ small, the buyer prefers non-commitment to commitment. The intuition is that the low type always gets zero, but the high type gets more rents when there is full separation. For δ close to zero, the high type gets $U(p|\overline{\theta}) + \delta U(0|\overline{\theta})$ instead of $(1+\delta)U(p|\overline{\theta})$.

5. The main difference between the two-type case and the continuum is that separation is possible but not usually optimal in the two-type case, while it is impossible in the continuous-type case. Intuitively, as $\Delta\theta$ becomes small, Type III contracts occur, requiring that both IC constraints bind. With more than two types, these IC constraints cannot be satisfied unless there is pooling almost everywhere.

3.3.4 Commitment with Renegotiation

We now consider commitment, but with the possibility that Pareto-improving renegotiation takes place between periods 1 and 2. The seminal work is Dewatripont's dissertation published in 1989. We will follow Laffont and Tirole's [1990] article, but in a nonlinear pricing framework.

The fundamental difference between non-commitment and commitment with renegotiation is that the "take-the-money-and-run" strategy of the low type is not possible in the latter. That is, a low-type agent that takes the high-type's contract can be forced to continue with the high-type's contract resulting in negative payoffs (even if it is renegotiated). Because of this, the low type's IC constraint is no longer problematic. Generally, full separation is possible even in a continuum, although it may not be optimal.

The Two-Type Case: We first examine the two-type case.

We assume that at the renegotiation stage, the principal makes all contract renegotiation offers in a take-it-or-leave-it fashion.[6] Given that parties have rational expectations, the principal can restrict attention to renegotiation-proof contracts.

[6]If one is prepared to restrict attention to strongly-renegotiation-proof contracts (contracts which do not have renegotiation as *any* equilibrium), this is without loss of generality as shown by Maskin and Tirole [1992].

It is straightforward to show that in any contract offer by the principal, it is also without loss of generality to restrict attention to a two-part menu of contracts, where each element of the menu specifies a first-period allocation, (q_1, t_1), and a second-period menu continuation menu, $\{(\underline{q}_2, \underline{t}_2), (\overline{q}_2, \overline{t}_2)\}$, conditional on the first-period choice, (q_1, t_1).

Renegotiation-proofness requires that for a given probability assessment of the high type following the first period choice, the solution to the program below is the continuation menu, where the expected continuation utilities of the high and low type under the continuation menu are $\{\overline{U}^o, 0\}$.[7]

$$\max \;\; \nu(\overline{\theta}\overline{q}_2 - \overline{U} - \frac{1}{2}\overline{q}_2^2) + (1-\nu)(\underline{\theta}\underline{q}_2 - \frac{1}{2}\underline{q}_2^2),$$

subject to

$$\overline{U} \geq \Delta\theta \underline{q}_2,$$
$$\overline{U} \geq \overline{U}^o,$$

where $\overline{U}$ is the utility of the high type in the solution to the above program. The two constraints are IC and interim-IR for the high-type, respectively.

The following partial characterization of the optimal contract, proven in Laffont-Tirole [1990], simplifies our problem considerably.

Theorem 32 The firm offers a menu of two allocations in the first period in which the low-type choose one for sure and the high-type randomizes between them with probability α on the low-type's contract. The second period continuation contracts are conditionally optimal given beliefs derived from Bayes' rule, $\underline{\nu} = \frac{\alpha p}{\alpha p + (1-p)} < p$.

We are thus in the case of Type I contracts discussed above with non-commitment. As a consequence, we know that $\overline{q}_1 = \overline{q}_2 = \overline{\theta}$ and $\underline{q}_2 = \underline{\theta} - \frac{\underline{\nu}}{1-\underline{\nu}}\Delta\theta$. The principal chooses $\underline{q}_1$ and α jointly to maximize profit. As before our results are ...

Results for Two-type Case:

1. $\underline{q}_1$ is chosen between the full-information and static-optimal levels. That is,

$$\underline{\theta} - \frac{p}{1-p}\Delta\theta \leq \underline{q}_1 \leq \underline{\theta}.$$

2. The probability of pooling is nondecreasing in the discount factor, δ. For δ sufficiently low, the full separation occurs ($\alpha = 0$).

3. As $\delta \to \infty$, $\alpha \to 1$, but for any finite δ the principal will choose $\alpha < 1$.

4. By using a long-term contract for the high-type and a short-term contract for he low-type, it is possible to generate the optimal contract above as a unique renegotiation equilibrium. See Laffont-Tirole [1990].

[7] We have normalized the continuation payoff for the low type to be $\underline{U}^o = 0$. This is without loss of generality, as first-period transfers can be adjusted accordingly.

Continuous Type Case: We look at the continuous type case briefly to note that full separation is now possible. The following contract is renegotiation-proof. Offer the optimal static contract for the first-period component and a sales contract for the second-period component (i.e., $q(\theta) = \theta$, and $t_2(\theta) \equiv C(\theta)$.) Because the second-period allocation is Pareto efficient in a first-best sense, it is necessarily renegotiation-proof. Additionally, no information discovered in the first period can be used against the agent in the second period (because the original contract guarantees them the maximal level of information rents) , so the first period allocation is incentive compatible. Without commitment, the principal cannot guarantee not to use the information against the agent.

Conclusion:

The main result to note is that commitment with renegotiation typically lies between the full-information contract and the non-commitment contract in terms of the principal's payoff. In the two-type case, this is clear as the lower IC constraint, which binds in the Type III non-commitment case, disappears in the commitment with renegotiation environment. In addition, the set of feasible contracts is enlarged in both the two-type and continuous-type cases.

3.3.5 General Remarks on the Renegotiation Literature

Intertemporal Price Discrimination: Following Hart and Tirole [1988] (and also Laffont and Tirole [1993, pp. 460-464], many of the above results can be applied to the case of intertemporal price discrimination.

Restrictions to Linear Contracts:

Freixas, Guesnerie, and Tirole [1985], consider the "ratchet effect" in non-commitment environments, but they restrict attention to two-part tariffs rather than nonlinear contracts. The idea is that after observing the output choice of the first period, the principal will offer a lower-rent tariff in the second-period. Their analysis yields similar insights in a far simpler manner. The nature of two-part tariffs effectively eliminates problems of "take-the-money-and-run" strategies and simplifies the mathematics of contract choice (a contract is just an intercept and a slope). The result is that the principal can only obtain more separation in the first period by offering more efficient contracts (higher-powered contracts). The optimal contract will induce pooling or semi-separating for some parameter values, and in these cases contracts are less distortionary in the first period.

Common Agency as a "Commitment" Device:

Restricting attention to linear contracts (as in Freixas, et al. [1985]), Olsen and Torsvik [1993], show how common agency can be a blessing in disguise. When two principals contract with the same agent and the agent's actions are complements, common agency has the effect of introducing greater distortions and larger rent extraction in the static setting. Within a dynamic setting, the agent's expected reduction of second-period rents from common agency reduces the high type's benefit of consuming the low type's bundle. It is therefore cheaper to get separation, and so the optimal contract has more information revealed in the first-period. Common agency effectively commits the principal to a second-period contract offer that lowers the high-types gain from lying. Martimort [1996] has also found a similar effect, but in a common agency setting with nonlinear contracts. Again, the existence of common agency lowers the bite of renegotiation.

Renegotiation as a Commitment Device vis-à-vis Third Parties:

Dewatripont [1988] studies a model in which in order to deter entry, a firm and its workers sign a contract providing for high severance pay (and therefore reducing the opportunity cost of the firm's production). Would-be entrants realize that the severance pay will induce the incumbent to maintain employment and output at high levels after entry has occurred, and therefore may deter entry. Nonetheless, there is an incentive for workers and the firm to renegotiate away the severance payments once entry has occurred, so normally this threat is not credible. But if asymmetric information exists, information may be revealed only slowly because of pooling, and so there is still some commitment value against the entrant (i.e. a third party). A related analysis is performed by Caillaud, Jullien and Picard [1995] in their study of agency contracts in a competitive environment (ala' Fershtman-Judd, [1987]) where two competing firms each contract with their own agents for output, but where secret renegotiation is possible. As in Dewatripont, they find that with asymmetric information between agents and principals, there is some pre-commitment effect.

Organizational Design as a Commitment Device against Renegotiation.

Dewatripont and Maskin [1995] consider the beneficial effects of designing institutions to prevent renegotiation. Decentralization of creditors may serve as a commitment device to cancel ex ante unprofitable projects at the renegotiation stage, but at the cost of some long-run profitable projects not being undertaken. In related work, Dewatripont and Maskin [1992] suggest that sometimes institutions should be developed in which the principal commits to less information so as to relax the renegotiation-proofness constraint.