Introduction to Repeated Games with Imperfect Public Monitoring

Jin-Ock Kim*

목 차

Introduction

In this review paper. we study repeated games in which players observe a public outcome which imperfectly signals the actions played. In this class of model. players cannot observe the other players' actions directly. but can observe imperfect and public signals about them. Obviously. the probability distribution over public outcomes depends on the players' actions. Thus. these are repeated games of moral hazard. Abreu. Pearce and Stacchetti (1990) and Fudenberg. Levine and Maskin (1994) have developed a beautiful and powerful set of solution concepts for these games. This is a rich class of problems with many significant economic implications. We illustrate a few examples in this class on games. The first example is the Cournot oligopoly (Green and Porter. 1984). in which firms sell output unobservably and the market price is a publicly observed outcome which is a stochastic function of total supply. The second example is the "noisy prisoner's dilemma" game. Players choose from their action set {Cooperation. Defection}. But. they do not observe their actions. but receive some noisy signal of their actions

* Associate Professor. Department of Economics. College of Economics and Commerce. Cheju National University

instead. The third example is the "partnership" game. There are two partners, $i=1,2$, each of whom has two unobservable actions, "work" or "shirk". The stochastic distribution of publicly observed outcome depends on action profile chosen by players. The fourth example is the game of "reputation" for quality. A single firm sets price p_t and chooses an effort e_t at some cost $c(e_t)$.

The publicly observed product quality is the stochastic function of effort level. Product quality is "High" with probability $p(e_t)$, where p is increasing in e_t. Consumers are willing to pay for high quality. The fifth example is the "Consumption smoothing and insurance (Green, 1987). There are a continuum of consumers, who each period get privately observed income shocks z_{it}. They then report their publicly observed incomes and make transfers among themselves. Transfers must be balanced. The final example is the self-enforced agency contracts (Levine, 2000). Each period t, the agent privately observes a cost parameter θ_t, and produces output y_t at cost $c(\theta_t, y_t)$. The output, but not the cost, is observed.

The Model

Consider the following stage game:

- There are n players. Each player $i=1,\dots,n$ chooses an action a_i from a finite set A_i. We call vector $a \in A \equiv \times_{i=1}^{n} A_i$ a profile of actions.
- Profile a induces a probability distribution over public outcomes $y \in Y$, where Y is a finite set. Let $\pi(y|a)$ the probability of y given a.
- Let $r_i(a_i, y)$ be player i's payoff if he plays a_i and the public outcome is y. Player i's expected payoff is defined as: $g_i(a) = \sum_{y \in Y} \pi(y|a) r_i(a_i, y)$.
- Player i's mixed strategy is $\alpha_i \in \Delta(A_i)$, where $\Delta(A_i)$ is the set of player i's possible mixed strategies. For a mixed strategy profile α, player i's expected payoff is defined as $g_i(a) = \sum_{y \in Y} \sum_{a \in A} \pi(y|a) \alpha(a) r_i(a_i, y)$.

In the Cournot oligopoly, a_i is firm i's quantity produced, while y is realized price. In the noisy Prisoner's dilemma game, a_i is player i's intended action, while y is actual action.

Following convention, let V be the convex hull of the set of feasible payoff vectors

$\{g(a)=(g_1(a),\ldots,g_n(a))|a\in A\}$. Let $\upsilon_i=\min_{\alpha_{-i}}\max_{a_i} g_i(a_i, \alpha_{-i})$ be player i's minimax value. The payoff vector v is individually rational if v_i is greater than υ_i for all player i. Let $V^*=\{v\in V|v_i\geq \upsilon_i$ for all $i\}$ be the set of feasible, individually rational payoffs. The Folk Theorem for the case of observable actions asserts that any individually rational payoff vector in V^* can be obtained by an equilibrium of the repeated game if the discount factor is sufficiently close to one.

Let's turn in detail to the repeated game, in which the stage game is repeated every period resulting in a public outcome y^t. The public history at the end of period t is $h^t=(y^0,\ldots,y^t)$. The private history of player i at the end of period t is $h_i^t=(a_i^0,\ldots,a_i^t)$. A strategy σ_i for player i is a sequence of functions $\{\sigma_i^t\}$, where σ_i^t is a function which maps each pair(h^{t-1}, h_i^{t-1}) to a probability distribution over A_i.

Definition 1 A public strategy for player i is a sequence of maps $\sigma_i^t : h^{t-1} \rightarrow \Delta(A_i)$.

We focus on public strategies because they are simple and lead to a nice structure for the game. Each public strategy profile $\sigma=(\sigma_1,\ldots,\sigma_n)$ generates a probability distribution over histories in the obvious way, and thus also generates a distribution over stage game payoff vectors. Players discount future payoffs with a common discount factor δ. Player i's objective in repeated games is to maximize the expected value of the discounted sum of her stage game payoffs. Player i's average discounted payoff for the game if she gets a sequence of payoffs $\{g_i^t\}$ is: $(1-\delta)\sum_{t=0}^{\infty} \delta^t g_{it}$.
This average discounted payoff is measured in the same units of stage-game poyoffs.

Definition 2 A public strategy profile $\sigma=(\sigma_1,\ldots,\sigma_n)$ is a perfect public equilibrium(PPE) if for each time t and public history h^{t-1}, the strategies form a Nash equilibrium from that point on.

This definition allows a player to contemplate deviating to a non-public strategy, but such deviations are not profitable for her because the other players stick to public strategies.[1] This means that a PPE is a perfect bayesian equilibrium (PBE) of the

1) With imperfect monitoring, there are no proper subgames since a player may be uncertain as to which of any information nodes he is at. So SPE might have not been applied. However, all possible nodes have the same distribution over opponents' play since opponents base their

repeated game since given opponents' sticking to public strategy, any private information he has is not payoff relevant. However, if some players are using their private informations, he might want to use a non-public strategy. This non-public strategy is not nearly well understood. Note that not all perfect bayesian equilibria are PPEs since PBE has not a recursive structure. A crucial point of PPE is that after any history h^t, a PPE induces a PPE in the remaining game. In repeated game with perfect observability of actions played, the set of subgame perfect equilibrium payoffs is stationary in the sense that it is the same starting from any period t. This story holds true with the set of PPE payoffs in repeated games with imperfect monitoring.

We look at the model of Green and Porter (1984) with public "trigger strategies:

1. Collusive phase: firms produce $q_1, \ldots, q_n$. If p_t is less than the trigger price p^*, then go the punishment phase 2.
2. Punishment phase: firms produce Cournot outputs $(q_1^c, \ldots, q_n^c)$ for T periods possible that $T = \infty$). Then return to collusive phase 1.

Above Green and Porter equilibrium is a PPE since strategies are public and play is Nash from every time forward.[2] Note that lower triggers imply less chance of punishment and more incentive to deviate, while longer punishment periods imply less incentive to deviate, but less of efficiency. This implies that firms can not achieve the first-best (monopoly profits) since there will be "price wars" in equilibrium.

Enforceable Actions and Self Decomposability.

Following Abreu, Pearce and Stacchetti (1990) and Fudenberg, Levine and Maskin (1994), we introduce a powerful set of techniques for characterizing perfect public

strategies on public information, not on private informations. So, there is no need to distinguish information nodes he is at.

2) In the punishment phase, it is optimal for a player to comply since opponent firms are playing a Nash equilibrim of the stage game. In the collusive phase, firms must find it optimal to play qc because increasing output increases single perid profits, but raises the probability of punishment being triggered.

equilibrium. In contrast to the Green and Porter analysis, we will think not in terms of strategies, but in terms of payoffs. The idea is that to "obtain" some individually rational payoffs(some $v \in V^*$), we should enforce certain actions at time t, which results in some public output to which we will attach continuation payoff from time t+1. We can think of this as an analogue to the principal-agent problem, where to induce the agent, the principal promised certain rewards and punishments in advance. The subtle difference here is that the promised rewards and punishments must correspond to payoffs in a PPE of the continuation game.

Definition 3: Let δ and $W \subseteq R^n$ be given. Profile α is enforcible with respest to W and δ if there exist $v \in R^n$ and a function w : Y → W such that, for all i,

(1) $v_i = (1-\delta)g_i(\alpha) + \delta \Sigma_y \pi(y|\alpha)w_i(y)$

(2) $\alpha_i \in \text{argmax}_{\alpha_i'} (1-\delta)g_i(\alpha_i', \alpha_{-i}) + \delta \Sigma_y \pi(y|\alpha_i', \alpha_{-i})w_i(y)$

where $w_i(y)$ is the ith component of w(y).

We say that the set $\{w(y)\}_{y \in Y}$ enforces α with respect to v and δ, and that v is decomposable with respect to α, W, and δ: condition (1) is that the target payoff v_i can be decomposed into current payoff $g_i(\alpha)$ and the expected continuation payoff ; condition (2) is essentially an incentive compatibility constraint. These conditions remind us of Bellman's equation for dynamic programming.

We let $B(W,\delta)$ be the set of all payoff vectors v such that for some α, (α, v) is enforceable with respect to fixed W and δ. If $W \subseteq B(W,\delta)$ for some δ, we say that W is self-decomposable. We also let $E(\delta)$ be the set of PPE payoffs.

Proposition 1: $E(\delta) = B(E(\delta),\delta)$

Proof. First we show $B(E(\delta),\delta)$ is the subset of $E(\delta)$. Suppose $v \in B(E(\delta),\delta)$. Pick α, w: Y → $E(\delta)$ such that w enforces (α, v). Now construct the following strategies. In period 0, play α. Then starting in period 1, play the perfect public equilibrium that gives payoffs $w(y_0)$. This is a PPE, so $v \in E(\delta)$.
Lastly we show that $E(\delta)$ is the subset of $B(E(\delta),\delta)$. suppose $v \in E(\delta)$. Then there

exists some PPE that gives payoffs v. Suppose in this PPE. play in period 0 is α. and continuation payoffs are $w(y_0) \in E(\delta)$. since continuation payoffs correspond to PPE play. The fact that no one wants to deviate means that (α, v) is enforced by $w : Y \rightarrow E(\delta)$. so $v \in B(E(\delta), \delta)$. Q.E.D.

Abreu. Pearce and Stacchetti (1986. 1990) call this factorization. The basic idea is that for any PPE. the corresponding payoffs can be decomposed or factored into current payoffs and continuation payoffs. All the continuation payoffs in a PPE correspond to PPE profiles. So those can be decomposed. and so on. They have a recursive structure.
Proposition 2. If W is self decomposable. then W is the subset of $E(\delta)$.

Proof. The hypothesis of self decomposability of W implies that any $v \in W$ belongs to $B(W, \delta)$. So there exist some $w : Y \rightarrow W$ and some α such that (α, v) is enforced by w. We construct a PPE that gives v. In period 0. play α. and for an realized outcome y_0. set $v_1 = w(y_0)$. Then $v_1 \in W \subset B(W, \delta)$. So again there is some α_1 and some $w_1 : Y \rightarrow W$ such that (α_1, v_1) is enforced by w_1. Continuing this way. we obtain the recommended strategies after each public history such that there are no profitable deviations. and which by construction give payoff v from time 0.

Above argument implies that $E(\delta)$ is the largest self-decomposable set.

Examples of Self Decomposability

To capture the main ideas of this review. We first consider the prisoner's dilemma game with perfect monitoring. The payoff matrix of this game is represented in table 1.

Table 1

	Cooperation	Defect
Cooperation	1,1	-1,2
Defect	2,-1	0,0

With perfect monitoring. $Y = \{(C,C),(C,D),(D,C),(D,D)\}$ where C and D are the short-hands for cooperation and defect respectively.

Claim If $\delta \geq 1/2$, the set $W = \{(0,0),(1,1)\}$ is self decomposable.

Proof. Our goal is show that $(0,0) \in B(W,\delta)$ and $(1,1) \in B(W,\delta)$ for $\delta \geq 1/2$. Consider (0,0) first. The action profile (D,D) and the payoff profile (0,0) are enforced by any δ and the function $w(y) = (0,0)$ since

$$0 = (1-\delta)g_i(D,D) + \delta w_i(D,D)$$

and for all $a_i \in \{C,D\}$,

$$0 \geq (1-\delta)g_i(a_i,D) + \delta w_i(a_i,D): \text{ incentive constraint}$$

Now we consider (1,1). The action profile (C,C) and payoff profile (1,1) are enforced by $\delta \geq 1/2$ and the function $w(C,C) = (1,1)$ and $w(y) = (0,0)$ for all $y \neq (C,C)$ since

$$1 = (1-\delta)g_i(C,C) + \delta w_i(C,C)$$

and all $a_i \in \{C,D\}$ with $\delta \geq 1/2$

$$1 \geq (1-\delta)g_i(a_i,C) + \delta w_i(a_i,C).$$

So W is self decomposable.

Now we consider a simple partnership game. There are two partners, $i = 1,2$, each of whom has two unobservable actions, "work"(w) and "shirk"(s). There are three output levels, $y = 12,8,$and 0 which depend on action profile chosen. If both players work the probability distribution over these levels is (1/3,1/2,1/6); If player 1 shirks and 2 works, the distribution is (1/3,0,2/3); if player 1 works and 2 shirks, it is (0,1/2,1/2); and if both players shirk, it is (0,0,1). Each player's ex-ante expected utility is half of expected output minus the disutility of his effort; work imposes a disutility of 3, and shirk cost 0. Action profiles (work, shirk) and (shirk, work) induce different probability distributions over outcomes.

Table 2 represents players' payoffs matrix depending on action profile chosen.

Table 2

	work	shirk
work	1,1	-1,2
shirk	2,-1	0,0

Let W be a small ball W in the interior of V^*. We claim that W is the subset of E(δ) for δ near enough 1. Then traditional Folk theorem extends to this example game. Proposition 2 says that if W is decomposable then W is the subset of E(δ). To prove the claim, we only show that W is self decomposable. We want to do this with above partnership game example. The strategy is that we show the boundary point v of W to be decomposable. A small ball W is divided into A, B, C, and D, as shown in Figure 1.

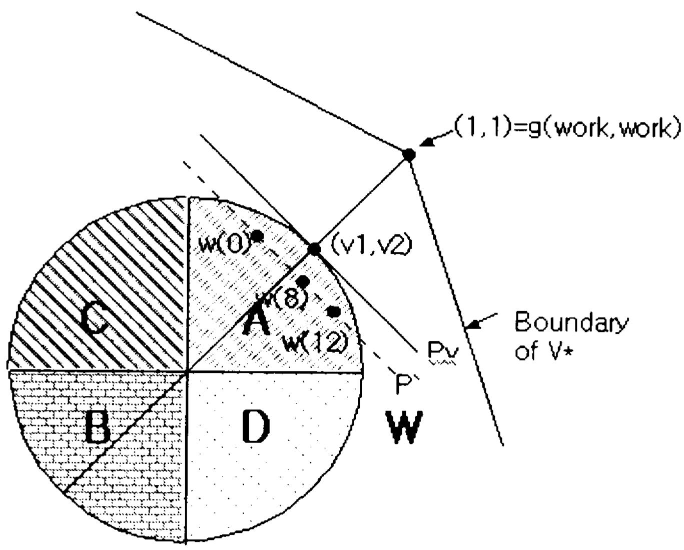

Figure 1

Let v be on the boundary on A. To decompose v, choose a=(work, work). To ensure that w(12), w(8), and w(0) belong to W, we select them to lie along a line P' parallel to the tangent hyperplane P_V to W at v (See Figure 1). Here P' takes the form $\{(v_1', v_2') | \beta_1 v_1' + \beta_2 v_2' = c\}$. To select w(12), w(8), and w(0), we solve the following system:

(i) $v_1 \geq 2(1-\delta) + \delta((1/3)w_1(12) + (2/3)w_1(0))$

(To prevent player 1 from shirking)

(ii) $v_2 \geq 2(1-\delta) + \delta((1/2)w_2(8) + (1/2)w_2(0))$

(To prevent player 2 from shirking)

(iii) $v_1 = 1-\delta + \delta((1/3)w_1(12) + (1/2)w_1(8) + (1/6)w_1(0))$

(player 1's expected payoff)

Above system implies that if δ is near enough 1, $w_1(12)$, $w_1(8)$, $w_1(0)$) will lie near the intersection of P' and the line through g(a) and v, and so will lie in W. Using the fact that continuation payoffs lie in P', we can replace each $w_2(y)$ in (2) by $(c-\beta_1 w_1(y))/\beta_2$. Finally we rewrite (i)-(iii) as equalities to obtain

$$\text{(iv)}\quad \begin{pmatrix} 1/3 & 0 & 2/3 \\ 0 & 1/2 & 1/2 \\ 1/3 & 1/2 & 1/6 \end{pmatrix} \begin{pmatrix} w_1(12) \\ w_1(8) \\ w_1(10) \end{pmatrix} \quad \begin{pmatrix} (v_1-2(1-\delta))/\delta \\ (\beta_2(1-\delta)-v_2)+c\delta \\ (v_1-1+\delta)/\delta \end{pmatrix}$$

Note that the matrix in (4) has full rank. So we solve the system to get $w_1(.)$.

Self-Decomposability and Extremal Equilibria

Abreu, Pearce and Stacchetti (1990) show that the structure of PPE often can be simplified by having the players use extremal rewards and punishments. The extremal points of a compact set $W \subset R^n$ are those points that aren't convex combinations of other points in W. Not all boundary points are necessarily extreme. All points in W are convex combinations of the extremal points.

Proposition 3. If (α, v) is enforceable with respect to δ and W, and W is compact, then (α, v) is enforceable with respect to the extremal points of W.

Proof(a brief sketch). By the hypothesis of the proposition, there exists a function w: Y

→ W that enforces (α,v). Then for any realized $y \in Y$, we can let w(y) be the convex combination of extreme points of W. Thus, we could simply replace w(y) with a randomization over extreme points that gives the same expected continuation payoff. Q.E.D.

These results imply that to achieve any equilibrium payoff in E(δ) we start with some profole α and then from the next period play a PPE that gives an extremal payoff. Likewise to achieve an extremal PPE payoff, we start with some profile α and then get another extremal PPE payoff(perhaps the same one). So this equilibrium will move between maximal rewards and punishments. In symmetric games, Abreu, Pearce and Stacchetti (1986) define strongly symmetric equilibria to be equilibria where players use symmetric strategies after every history. In this case, the set of PPE payoffs looks like Figure 2.

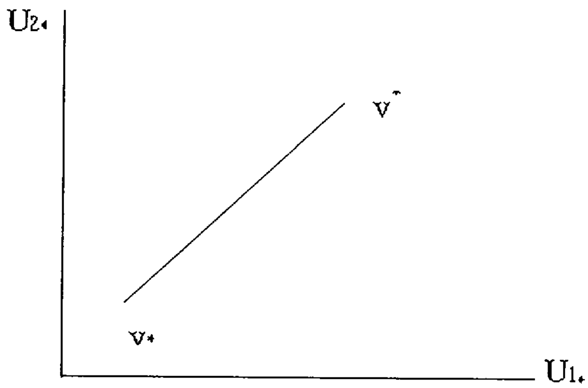

Figure 2

As seen in the above figure, there are only two extreme points in the set of symmetric PPE. So the extremal result is now very useful. It says that the equilibrium payoffs can be achieved by switching between v^* and v_*, giving a Green-Porter style equilibrium(here profile a^* corresponds to highest payoff v^*, profile a_* corresponds to lowest payoff v_*, profile a corresponds to intermediate payoff v):

- Phase I: Play a. Go to either phase II or III depending on y.
- Phase II: Play a^*. Go to either Phase II or III depending on y.
- Phase III: Play a_*. Go to either phase II or III depending y.

Like the Green-Porter equilibrium, strongly symmetric equilibria are generally inefficient. This is because the players can't do the optimal thing in every period: with positive probability they end up in Phase III doing something inefficient. Abreu, Pearce and Stacchetti (1990) go on to strengthen their extremal result by showing the following remarkable theorem. We present it without proof.

Proposition 4. Suppose Yis continuous and has constant support. If v is an extremal payoff of $E(\delta)$, then under mild regularity conditions, it must be enforced by extremal points of $E\delta)$.

The basic idea is that if v is extremal, then it can be thought of as the solution to some maximization problem. That is just to find the payoff that maximizes a weighted sum of the players' payoffs subject to incentive constraints. The choice variables are the present payoffs and the continuation payoffs. This turns out to be a linear programming problem with a convex choice set. These problems always have extremal solutions, and sometimes, all solutions are extremal.

The Folk Theorem

We have observed that Green-Porter strategies, or strongly symmetric strategies, lead to inefficient outcomes. Nevertheless, Fudenberg, Levine and Maskin (1994) show that this inefficiency arises because these strategies limit the space strategies, and go on to prove a version of the Folk Theorem. Fudenberg, Levine and Maskin's result requires two "identification" conditions.

(I) For all i, and α_{-i}, the $m_i = |A_i|$ vectors $\pi(y|a_i, \alpha_{-i})$ are linearly independent.

(II) For all i, j, there is some profile α such that the $|A_i| + |A_j|$ vectors $\pi(y|a_i, \alpha_{-i})$ and $\pi(y|a_j, \alpha_{-j})$ admit only one linear dependency.[3)]

3) Let $\Pi_i(\alpha_{-i}) = \pi(\cdot|\cdot, \alpha_{-i})$. We construct $(m_i + m_j) \times m$ matrix $\Pi_{ij}(\alpha)$ by stacking $\Pi_i(\alpha_{-i})$ on top of $\Pi_j(\alpha_{-j})$. Matrix $\Pi_{ij}(\alpha)$ has a linear dependence among its rows. This is because $\pi(\bullet|\alpha) = \sum_{a_i \subset A_i} \alpha_i(a_i)\pi(\cdot|a_i, \alpha_{-i}) = \sum_{a_j \subset A_j} \pi(\cdot|a_j, \alpha_{-j})$. In view of this linear dependence, has rank at most $m_i + m_j - 1$.

The first condition is that all actions of player i can be statistically distinguished. That is. player i's different actions do not induce the same probability distributions on outcomes. Furthermore. it requires that $|Y| \geq |A_i|$. which is arguably quite strong condition. The second condition says that if everyone is playing α. then can player i's actions be distinguished. and player j's actions being distinguished. but also their actions are distinguished form each other. The second condition is what we need to identify both a_i and a_j at the same time given some profile played by the others. $k \neq i, j$.

Definition 4. A set $W \subseteq R^n$ is smooth if (1) it is closed and convex: (2) it has a nonempty interior: and (3) at each boundary point v there is a unique tangent hyperplane P_v. which varies continuously with v.

We are in a position to present a version of Folk Theorem as follows.

Folk Theorem : Suppose dim $V=n$. and condition (I) and (II) hold. Then any smooth set $W \subset W$ in the interior of V^*. there exist some $\delta^* < 1$ such that for all $\delta \geq \delta^*$. $W \subset E(\delta)$.

The sketch of Proof. The proof has two steps. Step 1. FLM show that as $\delta \to 1$. then if it is possible to enforce some profile α today. then the pair (α, v) can always be enforced by choosing continuation payoffs that lie along some "tangent hyperplanes." or equivalently. along its translates. FLM then show that as $\delta \to 1$. we can fit a translate of a tangent hyperplane with essentially unbounded variation inside the boundary of the payoff set. This basically means that α can be enforced as $\delta \to 1$ if and only if one can find functions relating outcomes to continuation payoffs that enforce α and satisfy a particular linear equality.

Step 2. To finish the proof. FLM show that it is possible to find functions that relate outcomes to future payoffs. that (a) ensure that players will want to play α given these continuation payoffs. and (b) satisfy a particular linear equality. they show that conditions (1) and (2) ensure this. Q.E.D.

The Folk theorem applies to payoff vectors in the interior of V^*. which implies that we can't get exact efficiency with imperfect monitoring. The argument is simple and illustrative. Suppose that $\pi(\cdot|\alpha)$ has a support that is independent of α. And suppose that v is extremal but not a satic Nash equilibrium payoff. Because v is extremal, the only sequence of payoffs that gives average value v must have payoffs v in every period. So if PPE gives v, the first period strategies must specify a profile α such that $g(\alpha)=v$, and the continuation payoffs must be $w(y)=v$ regardless of any realized outcome. Note that continuation payoffs are independent of today's outcome. So unless α happens to be a static Nash equilibrium, someone will want to deviate.

References

Abreu, Dilip, David Pearce and Ennio Stacchetti (1986), "Optimal Cartel Monitoring with Imperfect Information,' Journal of Economic Theory, 39, 251-269.

Abreu, Dilip, David Pearce and Ennio Stacchetti (1990), "Toward of Theory of Discounted Repeated Games with Imperfect Monitoring," Econometrica, 58, 1041-1063.

Fudenberg, Drew, David Levine and Eric Maskin (1994), "The Folk Theorem with Imperfect Public Information," Econometrica, 62, 997-1039.

Green, Edward, and Robert Porter (1984), "Noncooperative Collusion under Imperfect Price Information," Econometrica, 52, 87-100.

Levin, Jonathan (2000), "Relational Incentive Contracts," American Economic Review, forthcomming.