Efficiency of Markets Under Moral Hazard With Side-Trading

Charles Kahn
Department of Economics
University of Illinois

Dilip Mookherjee
Statistical Institute
New Delhi

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign
Efficiency of Markets Under Moral Hazard With Side-Trading

Charles Kahn
Department of Economics
University of Illinois

Dilip Mookherjee
Statistical Institute
New Delhi
EFFICIENCY OF MARKETS UNDER MORAL HAZARD WITH SIDE-TRADING

by

Charles M. Kahn*

and

Dilip Mookherjee**

Current version February, 1991

*Department of Economics, University of Illinois Urbana-Champaign

**Indian Statistical Institute, New Delhi

We thank Andy Postlewaite, Herve Moulin and participants in the Summer Workshop at Northwestern University for useful comments. This research was supported by the National Science Foundation, through NSF grant SES-8821723.
This paper examines the efficiency of decentralized activity in market equilibrium under incentive constraints. We focus on insurance markets with moral hazard, under varying assumptions regarding the structure of information available to agents trading with one another. In particular, we examine contexts where, in addition to the unobservability of effort levels, trades between an insurer and insured are not observable to other insurance firms. Such contexts generate "incentive externalities," owing to the unenforceability of exclusive contracts involving limited insurance. We contrast this with contexts where trades and/or effort are publicly observable.

For each informational setting, we derive the corresponding notion of constrained-efficiency, by considering the problem of a social planner endowed with comparable amounts of information as private agents in the market setting. Decentralized activity in the market is represented by a general contracting game, with outcomes defined by a solution concept which allows coalitions of agents to coordinate their actions, subject to constraints imposed by information available about effort and trades. For each informational setting, we discover a close correspondence between market outcomes and the corresponding notion of constrained efficiency. In this sense, therefore, incentive externalities do not cause market outcomes to be constrained inefficient. However, they do imply a welfare gain from public observability of trades.
EFFICIENCY OF MARKETS UNDER MORAL HAZARD WITH SIDE-TRADING

INTRODUCTION

A prime question in welfare economics is the extent to which agents, although acting in a decentralized manner, can nevertheless achieve outcomes that are socially desirable. This paper investigates this question for economies with imperfect information -- specifically, in the context of markets with moral hazard. Such markets typically involve significant externalities between different firms that may potentially transact with a given customer. These externalities stem from the public good character of the customer's effort: the effect of varying the amount of trade offered by one firm affects the customer's effort, and thereby also the profits of other firms transacting with the same customer. This creates the need for exclusive contracts, where each customer trades with a single firm, and where the amount of trade is limited in order to generate requisite effort incentives. Such restrictions, however, create an incentive for customers to circumvent them, by engaging in additional trades with other firms. Exclusive contracts cannot be enforced if side-trades between the customer and other firms cannot be prevented. In such contexts, exchange is constrained not only by the unobservability of effort, but also by the possibility of side-trades with other firms.

This problem is not limited to insurance markets: the externalities from side-trades are likely to pose difficulties whenever hidden information or hidden actions are present. In credit markets, different lenders exert externalities upon each other, as the lending policies of one affect the probability that others will be repaid, by affecting the incentives of borrowers with respect to effort and project selection (Kletzer, 1984); see
The design of managerial compensation schemes is complicated by the ability of managers to purchase stock options and trade in futures markets on the side. Dynamic incentive schemes are also constrained if agents can borrow and lend (Rogerson (1985), Fudenberg, Holmstrom and Milgrom (1990)). In agrarian economies, the externality associated with moral hazard has been argued to constitute an important reason for the interlinking of credit and tenancy contracts (Braverman and Stiglitz (1982)). Similar externalities are exhibited in macroeconomic models based on labor contracts with imperfect information (Grossman, Hart and Maskin (1985), and Kahn and Mookherjee (1988)) and in centrally planned economies with black markets.

Previous literature, has, however, exhibited considerable disagreement regarding the implications of "incentive externalities" for the efficiency of decentralized market outcomes. On the one hand, papers by Pauly (1974), Helpman and Laffont (1975), Arnott and Stiglitz (1986) and Greenwald and Stiglitz (1986) -- as well as some of the papers cited above -- argue that with unobservable side-trades market outcomes are constrained-inefficient: specifically, there exist the scope for governmental tax-subsidy schemes to effect Pareto improvements. In contrast is the approach of Prescott and Townsend (1984), which extends the Arrow-Debreu theory of competitive equilibrium to moral hazard contexts. With an alternative notion of commodities and agents, viz. where consumers purchase (at given prices) units of (randomized) individually incentive compatible contracts over consumption-effort pairs, they succeed in extending the classical existence and welfare theorems of competitive equilibrium.

This disagreement motivates the following three basic questions: (i) What is the relevant model for competitive behavior in markets with moral hazard? (ii) What is the appropriate benchmark of efficiency for evaluating market
outcomes? (iii) How do market outcomes perform according to the appropriate notion of efficiency? The purpose of this paper is to develop a framework that addresses these questions. This framework will enable us to understand better the basis for arguments for the efficiency or otherwise of competitive markets with imperfect information. For the sake of analytical simplicity, we apply this framework to an example of an insurance economy with multiple potential insurance firms.

(i) What is the relevant model for competitive behavior in markets with moral hazard? The model we propose makes explicit the underlying assumptions regarding the structure of information, and the nature of trading and contracting processes. We study a Shubik-style trading game between firms and customers, under three alternative informational scenarios. In the first, effort levels, as well as all trades between every pair of agents are publicly observable. Consequently, a firm and a consumer can write contracts conditioned on the latter's effort, as well as on trades executed with other firms. In the second scenario, effort levels are unobservable, while trades continue to be observable. In this context, exclusive contracts can still be enforced, and are subject only to constraints imposed by the unobservability of effort levels. Finally, in the third scenario neither effort nor trades are observable; here transactions are constrained by the incentives of consumers to choose appropriate effort levels, as well as to trade with other firms.

The structure of observability of effort and trades defines the kinds of commitments that agents can enter into, and thereby the nature of contracts that can be enforced. We place no restrictions on the set of agents that can enter into a contract, thereby allowing the formation of arbitrary coalitions. These coalitions can coordinate their trading actions, subject to constraints imposed by opportunistic behavior of members with respect to
variables that are publicly unobservable. In the third informational scenario described above, for instance, a coalition of firms and consumers is constrained both by the unobservability of effort levels of consumers, as well as of their trades with different firms, both inside and outside the coalition. The ability of coalitions to form in order to "internalize externalities" is thus endogenously restricted in this fashion.

We argue that a natural solution concept for a game of multilateral contracting is an incentive constrained version of Aumann's Strong Equilibrium, i.e., where no coalition of agents should have an incentive to deviate with an incentive compatible contract between themselves. An important issue is the appropriate formulation of these incentive constraints. In order to capture the unobservability of side-trades, we impose (in addition to individual incentive constraints) the restriction that a deviating contract not be vulnerable to a further coordinated deviation by a pair of agents which is individually incentive compatible. The resulting solution concept is labelled Pairwise Incentive Compatible Strong Equilibrium (PICSE).

(ii) What is the appropriate benchmark of efficiency for evaluating market outcomes? The relevant efficiency criterion is typically represented by solutions to the problem faced by a hypothetical social planner who seeks to directly mandate actions of different economic agents in order to promote their welfares. The natural benchmark is to consider a social planner subject to exactly the same informational constraints that agents on the private market are subject to. We therefore consider the planning problem under three parallel informational scenarios. In the first-best context, the planner can observe effort levels as well as trades between every pair of agents, and can mandate levels of these variables throughout the economy. In the second-best scenario, all trades are observable, but efforts are not. And our new "third-best" context is one where the planner can verify neither trades nor
efforts. In the latter two contexts, the planner must issue instructions that are consistent with the incentives of agents to abide by them. Second-best allocations are thus constrained by the incentives of individual consumers to choose appropriate effort levels, while third-best allocations must incorporate additional coalition incentive compatibility constraints, relating to trades with private firms. The unobservability of trades typically generates a welfare loss, by requiring the provision of insurance to be restricted further than required by the need to provide effort incentives alone. A more detailed analysis of the third-best problem is provided in a related paper (Kahn-Mookherjee(1990a)).

(iii) How do market outcomes perform according to the appropriate notion of efficiency? For each of the three informational structures, we compare PICSE outcomes of the market game with solutions to the corresponding planner's problem. In the first-best and second-best scenarios, PICSE exist and there is an exact correspondence between PICSE outcomes and the set of constrained efficient outcomes. In the third-best context, we establish the existence of PICSE, as well as the Second Welfare Theorem: i.e., every third-best allocation can be achieved as the outcome of a PICSE, subsequent to initial lump-sum redistributions (We have also proved a limited version of the First Welfare Theorem: specifically, all PICSE outcomes in which the aggregate profit of firms is not too large, are third-best.)

Since for every informational context there exist constrained-efficient equilibria of the market game, our results appear to be at odds with those analyses that argue that competitive equilibria are necessarily constrained-inefficient. The reason is that such analyses employ a different notion of constrained-efficiency. Specifically, any assertion of inefficiency based on demonstrating the scope of tax/subsidy schemes to generate Pareto improvements, implicitly assumes that the planner has the ability to monitor
trades. In the absence of such monitoring, these tax/subsidy schemes cannot be implemented. Thus, arguments of this genre for the failure of a market with unobservable trades and effort implicitly use a second-best, rather than third-best efficiency standard. On the other hand, such arguments can be alternatively interpreted as a statement of the welfare gains from public monitoring of trades, an activity that some may argue is the natural province of governments.

The Prescott-Townsend approach, on the other hand, corresponds to a world where all trades are observable. This is implicit in their assumption that contracts defined over consumption-effort pairs (subject to effort incentive constraints alone) are enforceable. In other words, exclusive contracts are feasible, and there is no loss of generality in assuming that all consumers trade with a single centralized agency. Their demonstration of equivalence between competitive equilibria and efficient outcomes is mirrored in our model by the equivalence between PICSE outcomes and efficient outcomes in the second-best world.

The outline of the paper is as follows. Section 2 presents the basic model of insurance with moral hazard. Section 3 describes solutions to the planning problem under different informational constraints. Section 4 describes the general nature of the market game and the solution concept employed. Section 5 applies this to the insurance setting, and presents our main results concerning the relation between market and efficient outcomes.
A single risk-averse consumer wishes to purchase insurance from one or more risk neutral insurance companies. There are two states of nature: accident and no accident, and a single consumption good. An accident causes the consumer's endowment of this good to be reduced by $d$ units. By expending a utility-diminishing amount of effort the consumer can reduce the likelihood of an accident. The consumer's expected utility is described by:

$$W(x,e) = u(x_0) (1-p(e)) + u(x_1 - d) (p(e)) - e$$

where $x$ is a vector $(x_0, x_1)$ of net trades of state contingent consumption, and $u$ is a continuously differentiable, strictly concave, and strictly increasing function. Take state 0 to represent the no-accident state. The probability $p$ of an accident can be affected by the consumer's choice of a utility-decreasing effort denoted by the variable $e$. We assume that $e$ can take on two values: $e \in (0,1)$, and that

$$1 > p(0) > p(1) > 0.$$ 

Each insurer is risk neutral. If the consumer purchases policy $(x_0^i, x_1^i)$ from insurer $i$, and expends effort $e$, then the expected payoff to insurer $i$ is

$$V(x^i, e) = -x_0^i (1-p(e)) - x_1^i p(e).$$

There are a countable number of insurers. The total insurance received by the consumer is the sum of net trades accepted from firms he trades with:
\[ x = \sum_{i} x^i, \]

provided the sum converges. The effort level \( e \) is a public good, in the sense that all insurers' expected payoffs are affected by the level of \( e \) chosen by the consumer.

Let \( X \) be the set of **full insurance consumption levels** i.e.,

\[ X = \{ (x_0, x_1) \mid x_0 = x_1 - d \}. \]

Let \( Z(e) \) be the set of **zero profit policies** conditional on the choice by the consumer of effort level \( e \). That is,

\[ Z(e) = \{ x^i \mid V(x^i, e) = 0 \}. \]

\[ \text{------------------- INSERT FIGURE 1 HERE -------------------} \]

In Figure 1, net trades in the two states \( x_0 \) and \( x_1 \) are plotted on the axes. The origin represents the no-trade point. The straight lines \( Z(1) \) and \( Z(0) \) represent the trades at which firms would exactly break-even on aggregate, assuming that the effort level chosen is 1 and 0 respectively. The upward sloping line \( x_0 = x_1 - d \) represents the full insurance points. The dotted curved lines represent indifference curves of the consumer, assuming that the effort chosen is \( e = 1 \); the solid curved lines represent indifference curves corresponding to \( e = 0 \). The latter set of curves have a flatter slope, owing to the likelihood of an accident being lower with higher effort. For any given effort level, the indifference curves are tangent to the corresponding breakeven line along points of full insurance.
III. EFFICIENCY CRITERIA. THE SOCIAL PLANNING PROBLEM

Consider the problem of a social planner choosing an allocation for this economy. An allocation specifies trades \((x^i_0, x^i_1)\) between every firm \(i\) and the consumer, as well as an effort level \(e\) for the consumer. In the world where all trades as well as the consumer's effort are observable by the planner, the planner can choose any allocation whatsoever. Since we are interested in Pareto efficient allocations, we assume that the planner's objective is to maximize the welfare of the consumer, subject to the constraint that firms attain a predetermined aggregate expected profit \(\pi\). Then in the first-best outcome, the planner chooses any allocation with the property that \(x\), the consumer's net trade, solves

\[
\text{maximize } W(x, e) \quad \text{subject to } V(x, e) \geq \pi. \tag{i}
\]

In the second-best situation, trades are observable, but the consumer's effort choice is not. Hence the planner is constrained to allocations \(((x^i), e)\) satisfying the additional incentive constraint:

\[
e \in \arg\max_{\bar{e} \in (0,1)} W(x, \bar{e}) \tag{ii}
\]

In other words, once \(x\) is chosen, we cannot rely on the consumer to pick the efficient \(e\) to go along with it. Instead we must assume he will behave opportunistically. The optimal choice of \(x\) must be constrained by that opportunistic behavior. In general, for full or over-full insurance the consumer prefers low effort. When insurance is such that the payoff in the no-accident case is much greater than the payoff in the case of an accident,
the consumer is induced to exert effort. We assume that in the absence of any trade the consumer prefers to choose \( e = 1 \), that is:

\[
[u(0) - u(-d)][p(1) - p(0)] \geq 1. \tag{1}
\]

In Figure 1, the curve B represents the boundary between the region in which the consumer prefers \( e = 1 \) and the region in which the consumer prefers \( e = 0 \). We can therefore describe the consumer's preferences for pairs \( (x_0, x_1) \) in terms of the "reduced form" utility function

\[ Y(x) = \max (W(x,1), W(x,0)) \]

The heavily shaded line in the diagram represents one reduced form indifference curve.

------------- INSERT FIGURE 2 HERE --------------

Figure 2 describes the set of trades that enable firms to break even after incorporating the effect of that trade on the consumer's optimal effort choice. This is depicted by the heavy broken line, which coincides with \( Z(1) \) to the right of B, and with \( Z(0) \) to the left of B. Points to the southwest give positive profits in aggregate, the second best insurance policy corresponding to \( \pi = 0 \) is the utility maximizing point on the boundary for strictly risk averse consumers, the point will either be F, full insurance with zero effort, or S, less than full insurance with positive effort. At the latter solution, the consumer is restricted in the amount of insurance provided, to the minimal extent necessary to induce him to choose the high level of effort. The interesting case is the one where the second-best outcome for \( \pi = 0 \) is S.

Now consider what is likely to happen if the consumer can obtain additional insurance on the side from private firms, and the planner is
powerless to prevent such side-trades. Once the consumer has accepted the trade $S$, he will prefer to find an additional firm from whom to purchase additional insurance. Given the additional insurance, the consumer will choose to reduce effort. Nonetheless, the additional insurer will be able to offer an insurance policy that makes a profit even assuming $e = 0$, and the consumer finds this additional purchase desirable. They are not preferable (for the planner) to the second-best allocation, because in fact they make it no longer actuarially fair for the initial firm to supply $S$: given the reduction in effort, the initial firm will sustain a loss. Thus, the second best outcome may no longer be feasible if trades are unobservable.

If the planner also cannot monitor trades between the consumer and insurance firms, feasible allocations must incorporate the incentives of the consumer to enter into an unmonitored side-trade with some firms. A precise statement of the corresponding coalitional incentive compatibility constraint requires us to formulate the set of side-trades that third party firms will be willing to enter into. Such a set will certainly include any trade that is profitable irrespective of the consumer's effort choice. We will call these "safe side-trades." The set of safe side-trades is the cone

$$\mathcal{T} = \{ r \in \mathbb{R}^2 | V(r, 1) \geq 0 \text{ and } V(r, 0) \geq 0 \}$$

which is represented by the shaded area in Figure 2. Hence, a necessary condition for an allocation to be immune to an unmonitored side-trade is that $(x, e)$, the aggregate consumption and effort of the consumer, satisfy the following coalitional constraint:

$$(0, e) \in \arg\max_{r \in \mathcal{T}, e \in (0, 1)} W(x + r, \tilde{e}) \quad (iii)$$
If this constraint were not satisfied, some insurance firm could offer an additional trade which would be profitable and which would also make the consumer better off. However it is not clear that this condition is sufficient to rule out all side-trades, since (iii) considers only side-trades which make nonnegative profits irrespective of the consumer’s effort choice. It is conceivable that there exists a side-trade \( r \) and an associated effort level \( \tilde{e} \) which maximizes \( W(x+r,\tilde{e}) \), such that

\[
W(x+r,\tilde{e}) > W(x,e) \quad \text{and} \quad V(r,\tilde{e}) > 0
\]

and yet \((x,e)\) satisfies constraint (iii). In words, an insurance firm may offer \( r \) to the consumer, knowing that it will make positive profits as long as the consumer chooses his effort \( \tilde{e} \) optimally given the modified consumption \( x+r \). For the economy we examine, it turns out that (iii) is also sufficient to ensure that chosen allocations are immune to unmonitored side-trades. In other words, the planner’s optimum given constraints (i) - (iii) turns out to have the property that no pair of agents has any incentive to make any self-enforcing side-trades.\(^7\)

We thus define the third-best planning problem to be

Maximize \( W(x,e) \)

\[ x,e \]

subject to

\[
V(x,e) \geq \pi \quad (i)
\]

and \( (0,e) \in \arg\max r \in T,\tilde{e} \in (0,1) W(x+r,\tilde{e}) \quad (iii) \)

Note that constraint (iii) implies constraint (ii) above. We will call a solution to this problem a third-best outcome.\(^8\) The following properties of
third best outcomes are established in Kahn and Mookherjee (1990b)

**Proposition 1:** Given any preassigned level of profits $\pi$, a third best outcome exists. Any such third best outcome $(x_0, x_1)$ satisfies:

\[
\begin{align*}
    x_0 + \pi & \leq 0 \\
    x_1 + \pi & \geq 0 \\
    x_0 & \geq x_1 - d.
\end{align*}
\]

Thus, in particular, the set of third best outcomes lies within a bounded set. Subsequent to an initial redistribution to ensure attainment of the profit target $\pi$, a third best outcome does not give negative insurance and does not give more than full insurance. Such an outcome (for $\pi = 0$) is depicted in Figure 3. Let the point P be the maximum point on or below the actuarially fair locus $Z(l)$ such that the reduced form indifference curve through P lies entirely above the half-line extended from P parallel to $Z(0)$. If this point is north-west of the origin, then it is the third best outcome. It can be shown that a third-best outcome is either point F of the previous diagram (full insurance with zero effort), or it is an outcome involving high effort and strictly less insurance than is required to provide effort incentives (i.e., it lies to the right of the curve B). It is possible in a third-best outcome for firms to make expected profits greater than the pre-assigned level $\pi$; figure 4 gives an example where despite $\pi = 0$, they end up earning positive profits.

Our proposed coalitional constraint should be contrasted with a requirement that all agents have identical marginal rates of
substitution, sometimes called a no-retrade constraint. For our problem, the no-retrade constraint is a requirement that

\[ x \in X = \{(x_0, x_1) \mid x_0 - x_1 \geq d\} \]

That is to say, the no-retrade constraint would require that the consumer have full insurance, for only in this case would the marginal rate of substitution between income in the two states be identical for consumer and insurers. The no-retrade constraint has the same flavor as our third-best requirement, and has sometimes been employed as a restriction on contracts. (See, for example, Hammond (1987) for an analysis of this sort in a competitive environment without moral hazard). In the context at hand, the no-retrade constraint is an unsatisfactory formulation. Its plausibility is dependent on implicit, unstated assumptions about reactions of agents and it treats old and new insurers asymmetrically. For this constraint to make sense, new insurers must ignore the potential for the new insurance to alter the effort of the consumer, while old insurers take into account the incentive effects of any insurance they offer. In the diagrams the no-retrade constraint would always pick out the full-insurance, zero-effort point F.
The Contracting Game

In this section we formulate a contracting game that describes the process by which individual agents deal with one another in economies with private actions. We begin by describing the formulation in the insurance setting described above; at the end we briefly give a formal definition of a contracting game applicable to general economies with private actions.

In the insurance setting, the game will be played by many insurance firms and a single consumer. The consumer can potentially trade with multiple firms. It can enter agreements with firms individually or write a common contract with a conglomerate of firms. We define the notion of a contract more precisely below.

We visualize a two step procedure. In the second step, agents are engaged in a Shubik type trading process, where firms make offers of trades and the consumer makes requests to trade. If the request to a firm matches an offer by the firm the corresponding trade occurs; otherwise no trade occurs with that firm. The consumer also chooses the level of effort. Let \( x^i \) denote the trade between the consumer and firm \( i \) and \( x = \sum_i x^i \): then the payoffs for the consumer and firm \( i \) are respectively \( W(x, e) \) and \( V(x^i, e) \).

Formally, we assume a countable infinity of firms, \( i = 1, 2, \ldots \). Firm \( i \) offers the set \( S^i \subseteq \mathbb{R}^2 \). The consumer requests \( r^i \), \( i = 1, 2, \ldots \) and chooses effort \( e \). Trade occurs if \( r^i \in S^i \). We restrict the consumer to make a finite number of non-zero requests. This is our way of maintaining the free entry aspect of the model: it means there is always an unlimited number of inactive firms ready to make additional trades. \( N \) denotes the set of agents \( (0, 1, 2, \ldots) \), where agent 0 is the consumer.

Prior to the trading stage there is a communication/contracting stage.
Here different agents may attempt to coordinate their second-stage trading activities. If some of their second-stage activities are publicly observable, they may write contracts binding themselves to certain behavior. We assume these contracts over publicly observed variables are perfectly enforced by an outside agency.

The information structure for the economy indicates which second-stage actions are publicly observable. We deal with three different scenarios. In the full-information world, the enforcement agency can distinguish all second-stage actions; under moral hazard with observable trades, contract enforcers can observe all trade offers and requests, but not the effort chosen by the consumer. Finally if there is moral hazard and trades are unmonitorable, then contract enforcers can observe no actions whatsoever and first-period contracts do not bind agents at all. Formally, if $A_i$ is agent i's set of second-stage actions, then the information structure defines a partition $\theta_i$ of $A_i$, whose cells $\theta_i$ are publicly observable.

We now describe a contract more precisely. It is (1) a listing of the parties to the contract $I \subseteq N$ and (2) an enforceable commitment for each $i \in I$ -- that is, a specification of a cell $\theta_i$ in $\theta_i$ for each $i$. For a contract $c$, we let $I[c]$ denote the parties to the contract and $\theta_i[c]$ the commitment by player $i$ in the contract. As the information structure changes, so does the set of enforceable contracts.

In the first stage of the game, agents write contracts to coordinate their trades in the following way: Each agent $i$ proposes a contract $c_i$. A contract $c$ is executed if all parties to $c$ unanimously propose contract $c$, i.e. if $c_i = c$ for all $i \in I[c]$. Agent $i$ is bound to the commitment $\theta_i[c]$ only if $c$ is executed; otherwise he is bound to no commitment at all. Hence, as a function of first round contract proposals $\xi = \{c_0, c_1, c_2, \ldots\}$ agent $i$ is constrained to choose a second-stage action from the set
Thus a strategy for an agent in this two stage game consists of a contract proposal at the first stage plus a plan for second-stage actions $g_i$. This plan is a function of all contract proposals and must be consistent with the agent's commitments, if any, i.e.

$$g_i(c) \in F_i(c).$$

For this game we will consider three information structures: In the first-best information structure all actions by all agents are part of the information structure. In the second-best information structure, all offers to trade are part of the information structure, but the effort level $e$ is not. In the third-best information structure, no action is part of the information structure; thus the first stage of the game is redundant.

We conclude this section by describing the procedure for making any non-cooperative game into the second-step of a two-step contracting game. Take any non-cooperative game $\Gamma$ defined by strategy sets $A_i$ for $i \in N$ and payoffs

$$W_i : \bigvee A_i \rightarrow \mathbb{R}$$

for $i \in N$.

Take an information structure for this game. Let $\mathcal{C}$ be the set of all feasible contracts given the information structure; and let $\mathcal{C}^i$ denote the set of contracts to which $i$ is a party. Then the contracting game $\bar{\Gamma}$ is defined as follows: For each individual $i$ in the set of players $N$, the strategy set is

$$F_i(c) = \begin{cases} \theta_i | c_i \text{ if } c_i \text{ is executed} \\ A_i \text{ otherwise} \end{cases}$$
\[ D^i = \{ (c_i, g_i) \mid c_i \in \mathcal{C}^i \text{ and } g_i(\mathcal{C}) \in F_i(\mathcal{C}) \text{ for all } \mathcal{C} \in \mathcal{X} \mathcal{C}^i \} \]

Given a strategy vector

\[(\mathcal{C}, g) \in \bigwedge_{i \in \mathcal{N}} D^i\]

player \(i\)'s payoff is \(W_i(g(\mathcal{C}))\). 11

**Solution Concept**

A strictly non-cooperative formulation of \(\overline{F}\) would be unsatisfactory, because there would always be Nash equilibria where no contracts are signed and no trade takes place. (If all agents but one propose null contracts and non-participative trading strategies the final agent has no incentive to take any action either). Such equilibria appear unreasonable: if there are gains from trade or from writing contracts, we expect parties to coordinate their actions at the contracting and trading stages. For example, if the no-contract, no-trade equilibrium is in vogue, and we are in the full information world, a firm and a consumer have a mutual incentive to jointly propose a contract where they commit to exchanging first-best insurance while the consumer commits to first-best effort. This contract is enforceable under full information. Neither firm nor consumer could lose by proposing it, since if the other party fails to propose it, no one is bound by the commitment.

Coordinated strategies in \(\overline{F}\) seem natural provided such coordinated arrangements are (1) to mutual advantage and (2) enforceable. There are two ways that the arrangements can be enforceable: the publicly verifiable aspects can be turned into commitments which are publicly enforced; the unverifiable aspects must be credibly self-enforcing. For example, in a world where effort is unverifiable, a firm and consumer might try to write a
contract in which they commit to a trade which is profitable for the firm only if the consumer chooses high effort. Since the effort must be self-enforcing, the firm will only be willing to enter into such a contract if it were then in the consumer's own interest to choose high effort.

A natural way of modeling the formation of such arrangements is to formulate a notion of a credible deviation and then look for outcomes immune to credible deviations. That is what we proceed to do:

A coalition $C$ is a non-empty subset of $N$. An agreement is a pair $(a, C)$ where $C$ is a coalition and $a$ is a strategy vector in $T$. An agreement has the following interpretation: Taking as given the actions of those outside $C$, members of $C$ propose to play actions $(a_i)_{i \in C}$. We will say that agreement $(a, C)$ blocks strategy vector $b$ if

1. $a_i = b_i$ for $i \notin C$
2. $a$ is weakly preferred to $b$ by all $i$ in $C$, and strictly preferred by some $i$ in $C$.

Then a natural requirement for strategy vector $a$ to be a solution exhausting all possible gains from trade is that there be no agreement $(b, C)$ which blocks $a$. If so, $a$ is a Strong Equilibrium (SE) in the sense of Aumann (in contrast, a Nash equilibrium is a strategy vector which is blocked by no single-person agreement).

From our perspective, the difficulty with the SE is the credibility of the blocking agreement. For a strategy vector to be SE it must be immune to any unilateral deviations and to any joint deviation. In most situations, no SE exists; in particular as we will see, whenever individual incentive constraints bite there cannot be a SE.

It is more natural to restrict potential blocking agreements to a set of
"credible" agreements. We will define credibility of a deviation in terms of other agreements which block the deviation. We say that agreement \((a,C)\) blocks agreement \((b,D)\) if the agreement \((a,C)\) blocks the strategy vector \(b\) and \(D\) is a subset of \(C\). In other words, a blocking agreement is a coordinated deviation by a subset of members of the initial agreement which Pareto dominates the original strategy for the deviating group.

A minimal requirement for a credible deviating agreement \((b,C)\) is that every member of \(C\) have an incentive to stick with \(b\). Given an arbitrary agreement \((b,C)\) some member of \(C\) may prefer to deviate from the stipulated action, and this may jeopardize the welfare of other members of the coalition. In other words, we will want to require that blocking agreements be individually incentive compatible.

We call the agreement \((b,C)\) individually incentive compatible (IIC) if no individual in \(C\) has an incentive to deviate from his action in \(b\), given all other players' actions. If \((a,N)\) is IIC, then no player has an incentive to deviate from \(a\) -- in other words, \(a\) is a Nash equilibrium. If we take IIC to be the criterion for a credible multiplayer agreement, then we can use it to generate the corresponding equilibrium notion:

Strategy vector \(a\) is an Individually Incentive Compatible Strong Equilibrium (IICSE) if

1) \((a,N)\) is IIC

2) There does not exist an IIC agreement \((b,S)\) which blocks \(a\).

Note that every SE is an IICSE. The notion of IICSE is a natural solution concept for games in which individual incentive compatibility is a problem. In particular it is the analogue in non-cooperative settings of the Incentive Compatible Core. But as we will see, it is inadequate for
analysing games with side-trading, in which joint incentive problems arise. The difficulty is that IICSE does not go far enough in requiring deviations to be credible. An IICSE must be immune to all IIC joint deviations. The deviations themselves don't have to be immune to further joint deviations, only to individual deviations. In the case of insurance with unmonitorable trading, a joint deviation by insurance companies and the consumer may block a proposed allocation, but this deviation may be vulnerable to a further side-trade between some firm and the consumer. In a third best world, it is highly unlikely that a IICSE will exist, since any proposed allocation has to be immune to side-trading and also to alternatives which are themselves vulnerable to side-trading.

These considerations motivate the introduction of a further restriction on credible deviations and an associated equilibrium concept, the Pairwise Incentive Compatible Strong Equilibrium:

**Definition:** Agreement \((b,S)\) is **Pairwise Incentive Compatible (PIC)** if 1) \((b,S)\) is IIC and 2) provided \(S\) has more than two members, there does not exist an IIC two-player agreement which blocks \((b,S)\).

Note therefore that every IIC one- or two-player agreement is PIC.

**Definition:** \(a\) is a **Pairwise Incentive Compatible Strong Equilibrium (PICSE)** if

1) \((a,N)\) is PIC.

2) There does not exist any PIC agreement \((b,S)\) which blocks \(a\).

It should be readily apparent that this is the natural generalization of IICSE. The substantive difference is the requirement that joint deviations be immune to further deviations by pairs from the deviating coalition.
Proposition: If strategy vector $a$ is IICSE, then it is PICSE.

Proof: By condition 2) of IICSE, $a$ is not blocked by any two-player IIC agreement. Combined with condition 1) this demonstrates that $a$ is PIC. If there is no IIC agreement which blocks $a$ (condition 2 again) there is no PIC agreement which blocks $a$.

Note that if $a$ is not IIC then it is not PIC; if $a$ is not PIC, then $a$ is not PICSE, ICSE or SE. Thus all of these equilibria are refinements of Nash equilibrium. We can continue iteratively, defining Trio-wise Incentive Compatibility, Quad-wise Incentive Compatibility and so forth. The definition of N-Wise Incentive Compatible Strong Equilibrium would turn out to be essentially the Coalition Proof Nash Equilibrium of Bernheim Peleg and Whinston.\textsuperscript{13}

Nonetheless, for the problem at hand, we do not need to go to the complete formulation; we get identical results to CPNE by sticking with the easier notion of PICSE. The reason we can stop with PICSE is that all benefits of side-trading can be achieved within two-player coalitions.

v. APPLICATION TO THE INSURANCE MARKET

First- and Second-Best Economies

We now use these equilibrium concepts to analyse the insurance market game in the various informational contexts we have described.

Proposition 2: In the full information world, SE outcomes, IICSE outcomes and PICSE outcomes are identical, and coincide with the first best outcomes.
Proof. Clearly an allocation which gives negative profit to any firm is not individually incentive compatible: That firm would block by refusing to trade. Thus such an allocation cannot be PICSE, ICSE or SE.

Consider any outcome giving less than first best utility. By assumption, there exists a passive firm earning zero profits. Form the coalition of consumer and firm, where they exchange the first best exclusive contract. (In other words, the firm and the consumer specify that the firm will offer the first best level of insurance, the consumer will request the first best level of insurance and engage in the first best level of effort and request zero insurance from any other firm. Since all these actions are contractable, the contract, if consummated, completely restricts second-stage actions. If the contract is not consummated, second-stage actions are immaterial; assume both parties engage in zero trade in that event.) Since this two-agent blocking agreement is individually incentive compatible, the initial allocation could not have been PIC, therefore it cannot be PICSE, ICSE or SE.

In an allocation which gives the consumer first best utility, firms earn zero profits, so an allocation with positive profits must give a consumer less than first best utility. From the above argument such an allocation cannot be PICSE. We conclude that in every PICSE, all firms receive zero profits and the consumer receives first best utility.

It remains simply to show that SE exist, and yield the first best outcome, for then the SE must be PICSE and ICSE as well. We claim the following is an SE. Firm 1 offers the first best exclusive contract to the consumer who request the same. The consumer offers first best effort and no contract with any other firm. The associated second-step actions are as described above. All other firms offer nothing and do not trade in the second step. Suppose there exists a block (b,S). The outcome must Pareto dominate
the first best outcome, a contradiction

**Proposition 3:** If effort is unobservable but trades are observable, the IICSE outcomes and the PICSE outcomes are identical, and coincide with the second best outcomes.

**Proof:** Analogous to the above.

Note that if the incentive constraint in the second best problem bites, there is no SE in a second-best world.

**The Third-Best Economy**

Finally we turn to an examination of the third-best world. The analysis requires the set of agreements to be complete in a particular sense. By ensuring that player 0's strategy space is complete we avoids paradoxes that arise when a player has a continuum of desirable deviations but no "best" deviation. The modification we describe below to complete the set is a special case of the general procedure described in Kahn-Mookherjee (1991).

We complete the space of agreements by extending the strategies available to player 0 to include all "limits," (that is to say, all sequences) of strategies in its initial strategy space. Let \( A^* \) be the set of strategy vectors in which player 0 is allowed to play a limit strategy. We define payoff for any element of \( A^* \) to be the limit of the infimum along the sequence for player 0, and to be the infimum for any other player. Let \( \mathcal{A}^* \) be the set of agreements with this extended strategy space. The definitions of IICSE and PICSE now take agreements to be in the set \( \mathcal{A}^* \). The theorems of the previous section are unaffected by using this extended set of agreements (although some proofs become more complicated).
Proposition 4: When neither trades nor effort is observable, a PICSE outcome \((x,e)\) satisfies

\[
\begin{align*}
(i) \quad W(x,e) & \geq \max_e W(0,e') \\
(ii) \quad V(x^i,e) & \geq 0 \quad \text{for all } i \\
(iii) \quad W(x,e) & \geq \max_e W(x,e') \\
(iv) \quad W(x,e) & \leq W_{SB}
\end{align*}
\]

where \(W_{SB}\) is the insured's expected utility in the second-best outcome.

Proof: If any of (i) - (iii) are not satisfied, then there exists a singleton deviation which blocks. If it is not IIC, it cannot be an equilibrium.

(iv): use (ii) and (iii) to conclude \((x,e)\) is feasible in the Second Best problem.

Proposition 5: Any PICSE outcome \((x,e)\) is feasible in the third-best problem with zero profits.

Proof: From proposition 2, \(V(x,e) \geq 0\). Suppose the other condition of third best feasibility does not hold -- i.e. there exists \(r \in T\) and \(e' \in e\) such that \(W(x+r,e') > W(x,e)\). Consider the following two-agent agreement. A passive firm \(j\) offers \(r\), the consumer chooses \(e\), requests \(r\) from that passive firm and \(x\) from all the others. If this is IIC we are done. If not then the consumer must find it profitable to deviate. Suppose he has an optimal deviation. Then that optimal deviation plus the offer of \(r\) by the passive insurance company forms a two player IIC agreement which blocks, a contradiction. If there is no best deviation for the consumer, then any sequence of deviations the utility of which approaches the supremum serves as the consumer's strategy in the two-player IIC agreement.
and the same conclusion follows.

**Proposition 6**: Every third best outcome can be achieved as a PICSE with lump sum redistribution.

**Proof**: Start with the case where in the third best outcome aggregate profits are equal to 0. If profits in the outcome are greater than zero, redistribute those profits and repeat the argument.

Consider the following strategies: Player $i$ offers the third best contract, players $j = 2, 3 \ldots$ offer the safe cone, the consumer requests the third best contract from $i$, nothing from the other players, and chooses the third best level of effort.

First we show that there is no coalition which blocks the strategy and gives firm $i$ non-negative profits. For if there were, the outcome would Pareto dominate, and therefore must violate the third-best constraint. Since the coalition excludes some firm the purchase of $r$ is available. So the blocking coalition is not IIC.

Next we show that there is no coalition which blocks the strategy and gives firm $i$ losses. To give firm $i$ losses means that the consumer continues to buy the third best level $x^*$ from him and switches from a third best effort of $i$ to a deviating effort of 0. If the new total purchase $x$ induces effort 0 and $x - x^*$ makes non-negative profits at an effort level of zero, then $x - x^*$ makes non-negative profits at an effort level of one as well. Thus the total purchase from the other firms is an element of the safe cone, a contradiction.

From the above we conclude that the strategy vector 1) is PIC and 2) is not blocked by any PIC agreement. Finally, by the previous theorem, any outcome which Pareto dominates the third best problem does not satisfy these two conditions. Thus this strategy vector is a PICSE.
The final theorem shows that no stronger equilibrium notion will work in this problem as long as the third best constraint is binding:

**Proposition 7:** When neither trades nor effort are observable, there exists no IICSE if the third-best outcome is Pareto dominated by a second-best outcome.

**Proof:** Suppose a is IICSE. By proposition 1, it is also PICSE. By proposition 3 it is feasible in the third best problem, so consumer utility cannot be greater than the third best level at the given profits. Then the grand coalition could move to the second best allocation, which is an IIC block. Contradiction.

**Discussion**

Roughly speaking, the equilibrium in the third-best case works as follows: firm 1 offers the third-best contract, while all other firms offer the set $T$ of safe trades. The consumer accepts the third-best trade from firm 1, doesn't request any trade from other firms, and chooses the third-best effort. 14

It can happen that the firm offering the third-best contract will receive positive profits; nonetheless the identical competing firms are unable to undercut the offer. This is because any competing firm knows that additional offers will push the consumer to a point where he prefers to alter his level of effort. Thus in any equilibrium, the active firm has a first mover advantage.

Since each firm is identical, the equilibrium chooses which firms get to offer how much of the third best contract, and in different equilibria different firms receive different payoffs. Given the total distribution of others' assignments, there is no other assignment that any insurer prefers. 15
As the last theorem demonstrates, there is no strong equilibrium as long as the third-best outcome does not lie on the second-best frontier. The problem is that the coalition of the whole can block with a deviation which is individually incentive compatible, namely a second-best outcome. Of course this deviation is itself vulnerable to a deviation involving a coalition of the consumer and some firms, but such coalitional incentive constraints are not imposed on the blocking deviations in an IICSE. Thus it is precisely the lack of coalitional constraints on allowable deviations that precludes the existence of SNE and IICSE, while their inclusion allows the existence of PICSE.

VI. SUMMARY

We have examined an economic environment with moral hazard and the possibility of unmonitored trading. We modeled the situation as a market game and compared the equilibria under varying degrees of power to contract, corresponding to varying degrees of observability of actions. Our main result is that Pairwise INcentive Compatible Strong Equilibrium achieves the constrained optimal allocation with the appropriate level of incentive constraint. In the case of unmonitorable trading we proposed a new notion: third best contract which maintains this relationship.

Of the three different information structures examined, it is structure (b), where individual trades and consumption levels can be publicly verified, which most closely resembles the Prescott-Townsend formulation. The alternative "market-failure" approach models the outcome of decentralized markets with information structure (c), in which individual trades cannot be monitored. This suggests the following interpretation of the discrepancy between the two approaches:
On the one hand, the market failure approach involves an implicit assumption that governments have access to better monitoring technology than do private firms. Thus it is unsurprising that the market failure approach concludes that governments have a role in remediating externalities arising from moral hazard. Nonetheless, the assumption may be natural in some circumstances: the power of a government to tax a trade implies the ability of the government to monitor that trade. If governments are relatively efficient at monitoring trades with third parties (e.g., because such information has public good characteristics) then governments may in fact have a natural corrective role.

On the other hand, the Prescott-Townsend approach implicitly assumes extensive capability to monitor agents' trades. In the absence of moral hazard agents do not need to monitor each other's trades in the corresponding Arrow-Debreu competitive equilibrium; but in moral hazard settings, as we have seen, such monitoring is crucial. The fact that the equilibrium we derive is different from those proposed in previous analyses, and the fact that our equilibrium does have some efficiency properties, do not erase the conclusion that there is a welfare loss from the inability to monitor trades.

The similarity of the formalism of the Arrow-Debreu formulation in situations with and without moral hazard conceals the difference in importance of monitoring in the two settings. This suggests the need for caution in choosing the appropriate notion of commodities and agents in the presence of imperfect information. Indeed, one may wonder whether an allocation mechanism such as used by Prescott-Townsend may legitimately be described as decentralized, since it requires such extensive monitoring. If in fact we view the degree of monitoring as a defining characteristic of the degree of centralization of a resource allocation mechanism, our analysis demonstrates the need for centralized mechanisms in the presence of moral hazard.
Our analysis also provides a critical perspective on the Coase (1960) theorem, complementing the analyses recently presented by Chari and Jones (1988) and Mailath and Postlewaite (1988). Their critiques are based on the problems created by the presence of private information; we focus on the instability of coalition formation. In the absence of sufficient monitoring of side-trades, our analysis identifies a form of collusive free-riding by subcoalitions as an impediment to efficiency-improving contracts. Interestingly, the severity of this free-riding increases as the coalitions grow in size and the number of possible deviating subcoalitions increases. Thus the analysis provides a formal justification for models limiting cooperation to small coalitions, and a potential insight into problems faced in generating desirable cooperative outcomes for large coalitions in a wide range of economic environments.

We believe this approach provides insight on the role of governments in enhancing the efficiency of decentralized mechanisms. Do governments in fact have a comparative advantage in monitoring private actions? If so, are taxes the natural way to exploit this advantage in the public interest? Alternatively, is there anything to choose between the role of governments in monitoring private actions and enforcing contracts written between private parties, compared to more centralized allocation mechanisms involving taxation? It will be worthwhile to extend the examination of monitoring to contexts other than moral hazard in which "market failure" is claimed to justify government intervention — for example, private monopoly or common resource externalities.
REFERENCES


__________ and ____________, "Basic Analytics of Moral Hazard," mimeo. 1984a.


__________ and ____________, "Equilibrium in Competitive Insurance Markets, II: Existence and Nature of Equilibrium"


Hammond, Peter "Markets as Constraints: Multilateral Incentive


Mailath, George and Andrew Postlewaite, "Workers versus Firms: Bargaining over a Firm's Value," working paper, department of economics, University of
Pennsylvania, June 1988


This has a close relation with the notion of a Coalition Proof Nash Equilibrium, proposed by Bernheim, Peleg and Whinston (1986); see also Kahn Mookherjee (1991). It turns out that for the game in hand, the outcomes generated by the two concepts coincide. Intuitively, this is because most relevant deviations involve only pairs comprising one consumer and one firm. For a formal analysis of the Coalition Proof Nash Equilibria of the game with unobservable trades and effort, see Kahn and Mookherjee (1990).

In a sense, the choice of insurance as the application for our problem, while dictated by the history of this literature, is unfortunate. In the formal specification of the game in the previous section, we distinguish between contracts and spot actions. In this application, spot trades are themselves insurance policies. In cases where the ability to monitor is limited, we will nonetheless assume that the consumer is able to collect payments from the insurance company in the case of an accident and the insurance company is able to collect payments from the consumer in the case of no accident. If this strains the reader's imagination, he may prefer to consider $x_1$ and $x_2$ as concrete goods which can be traded spot (on informal markets, for example).

This framework has been considered by numerous authors, among them, Arnott and Stiglitz [1982, 1987], Pauly [1974], Hellwig [1983], and in related formulations, Helpman and Laffont [1975]. We will closely follow the basic analytics set out by Arnott and Stiglitz [1982].
More generally, the planner will incorporate a minimum expected profit target for each firm separately. It is straightforward to check that the more general problem is equivalent to (i), since the former is solved by solving (i) as a first step, and then distributing trades among firms to attain the profit targets for each of them separately. We shall continue to use the same aggregate formulation in the second and third-best scenarios as well, for the same reason.

The reader may check that the indifference curve of the consumer (assuming e=0) which passes through \(S\) has slope less than \((1-p(0))/p(0)\), which is the slope of \(Z_0\).


A formal proof follows an argument analogous to one which will be used in Section 5. Intuitively, any insurance firm \(i\) offering a side-trade to the consumer has to worry not only about the effort chosen by the consumer in the event of having consumption \(x+\hat{r}\). It has to consider the possibility that given the additional trade \(\hat{r}\) offered by \(i\), the consumer may have an incentive to enter into yet another side-trade \(\tilde{r}\) with a different insurance firm, and the effort level \(e\) that is optimal for the consumer given consumption \(\hat{x}+\hat{r}\) may cause the provider of \(\hat{r}\) to lose money. A more detailed discussion of the planning problem with unobservable side trades is pursued in Kahn and Mookherjee (1990b).

When more than one allocation solves this problem, we will restrict attention to those maximizing firm profits \(V(x,e)\).
Note that even though the supplier of insurance is making positive profits, the planner cannot increase the consumer's utility by redistributing consumption in his favor; the resultant allocation would violate the coalitional constraint (iii). That is, although the new insurance would give the consumer higher utility, it would also cause him to want to purchase still more insurance, which other firms would be willing to supply at prices which are profitable irrespective of the consumer's effort. The additional insurance would cause the consumer to switch his effort choice in a manner which would make existing insurance providers lose money.

This structure is crucial to our model, because it allows agent i to offer contracts which commit himself to trades conditional on agreements by other agents committing themselves to "the other side of the bargain." Without this quid pro quo feature, other parties could take advantage of unconditional offers by agent i, while refusing to bind themselves. For example, in the insurance setting with observable effort and trades (the first-best world) an insurance firm offering a first-best contract could be exploited by an insured individual who accepts the perfect insurance offered, while refusing to bind himself to first-best levels of effort.

Despite the fact that the contracting game is a two-stage game, we do not impose any perfection requirements: Our formulation does not require action choices to form an equilibrium in the trading game subsequent to off-equilibrium contract proposals. For the case which we are examining it appears immaterial.

See Boyd-Prescott, Marimon, James Kahn, and Berliant.

See Kahn-Mookherjee (1991) for more details.
The equilibrium therefore requires that firms offer an infinite number of alternative trades which the consumer does not accept. It captures the notion of a world with free entry, where potential entrants stand ready to enter and offer safe trades. The importance of such "threat contracts" is noted by Arnott and Stiglitz (1987).

Andy Postlewaite has suggested an extension of our model where the consumer visits different firms in sequence. Each firm would worry about the incentive of the consumer to trade with firms he will subsequently visit, and implications for the effort choice. In this setting there may be a unique CPNE satisfying suitable perfection constraints, where the first firm that the consumer visits obtains a first mover advantage.

See Myerson (1988) for a similar view in the context of adverse selection economies.