FACULTY WORKING PAPER 92-0101

Optimal Contracts With Costly State Verification: The Multilateral Case

Stefan Krasa
Department of Economics
University of Illinois

Anne P. Villamil
Department of Economics
University of Illinois

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 92-0101

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

January 1992

Optimal Contracts With Costly State Verification: The Multilateral Case

Stefan Krasa
Anne P. Villamil

Department of Economics

Optimal Contracts With Costly State Verification: The Multilateral Case

Stefan Krasa   Anne P. Villamil*

First Draft: January 1991   This Draft: December 1991

Abstract

The purpose of this paper is to derive the structure of optimal multilateral contracts in a costly state verification model with multiple agents who may be risk averse and need not be identical. We consider two different verification technology specifications. When the verification technology is deterministic, we show that the optimal contract is a multilateral debt contract in the sense that the monitoring set is a lower interval. When the verification technology is stochastic, we show that transfers and monitoring probabilities are decreasing functions of wealth. The key economic problem in this environment is that optimal contracts are *interdependent*. We are able to resolve this externality problem using abstract measure theoretic tools.

---

*Address of the authors: Department of Economics, University of Illinois, 1206 South Sixth Street, Champaign, IL 61820

We wish to thank Mark Feldman, Wayne Shafer and Nicholas Yannelis for useful comments. We also gratefully acknowledge financial support from the National Science Foundation (SES 89-09242).

1 Introduction

In the Arrow-Debreu model complete insurance markets exist and agents are able to attain unconstrained Pareto efficient consumption allocations. In addition, the structure of the set of financial securities that support these allocations is indeterminate (i.e., the Modigliani-Miller Theorem states that a firm's debt-equity ratio is irrelevant when there are no market imperfections). Casual observation and systematic study, however, suggest that firms have determinate debt-equity ratios and that insurance markets are incomplete. The costly state verification model, proposed by Townsend (1979), provides one plausible explanation of these outcomes that is consistent with (constrained) Pareto efficient behavior. In particular, Townsend introduces an information friction into the Arrow-Debreu model with two essential elements. First, agents have asymmetric information: All agents know the (common) distribution of random variables (i.e., endowments) in the economy, but the realization of a particular agent's random variable is costlessly observed only by the agent himself. Second, a technology exists that can be used to publicly announce the realization to all agents ex post, but it is costly to use. The model has proved useful for analyzing economic problems with both deterministic and stochastic verification technologies.

When costly state verification is deterministic (i.e., monitoring occurs with either probability one or zero), Townsend proves that the optimal contract that supports (information and resource) constrained Pareto efficient consumption allocations resembles debt because the monitoring set is a "lower interval." That is, it is optimal to monitor all announcements below a certain cut-off point, and these outcomes are interpreted as states of costly (but efficient) bankruptcy. All announcements above the cutoff point are not monitored, and these outcomes are interpreted as states of solvency. The lower interval result establishing the optimality of debt is important because it is consistent with many stylized facts observed in actual markets.[1] For example, it is consistent with the prevalent issue of debt by firms, its payment

[1]Lower interval results are also obtained by Gale and Hellwig (1985), but they assume that only the agent who pays for monitoring gets the information. The distinction between public and private monitoring does not matter in a two agent economy, but it is important when there are many agents. See Williamson (1986) and Krasa and Villamil (1991a) for multiple agent costly state verification environments with private monitoring reports.

characteristics, and key institutional features of bankruptcy procedures.[2]

Unfortunately, existing lower interval results have been established only under several restrictive assumptions: Agents are either assumed to be risk neutral or their trades are exogenously restricted to be symmetric, separable in endowments, and bilateral.[3] Townsend (1979, p. 281) notes that these restrictions are "unpleasant" because they are motivated by technical, rather than economic considerations. Further, they may preclude optimal risk sharing arrangements even in two-agent contracting problems.[4] More fundamentally, however, Boyd and Prescott (1987) argue that coalitional structures are important for understanding many economic phenomena.[5] For example, Boyd and Prescott (1986) study explicitly *multilateral* contracts which (p. 217) "condition the consumption allocations of [agents] on *group* experience as well as on observables for individual(s)." Their model (with adverse selection and two agent types) gives rise to welfare-improving financial intermediary coalitions which exhibit key characteristics displayed by actual intermediaries. They (p. 231) note: "An extension which is not so easy [in their model] is to allow for more than two agent or project types." Such an extension may prove to be straightforward in our framework (cf., Section 5).

The second class of problems that the costly state verification model has proved useful for analyzing are environments where the verification technology is stochastic (i.e., monitoring need not occur with probability one). In this case the optimal contract that supports (information and resource) constrained Pareto efficient consumption allocations specifies transfer and monitoring procedures that resemble those commonly used by insurance companies and tax revenue collection agencies: The optimal contract has transfers and monitoring probabilities that are monotonically decreasing functions of an agent's reported wealth. Border and Sobel (1987) prove this result in a model with stochastic monitoring, two risk neutral agents (one having a

---

[2]We discuss relevant aspects of US bankruptcy procedures in Section 5.

[3]Recently, Winton (1991) considers the case where all but one agent (say the firm) have a riskless endowment. He shows that contracts which resemble subordinated debt are optimal in this framework.

[4]Townsend (1979, p. 281) provides an example where two agents have utility functions of the form $u(c) = c^{\alpha+1}/(\alpha+1)$, where $-1 < \alpha < 0$. The optimal symmetric transfer function implied by this common utility specification is not separable in endowments as required by the exogenous restriction.

[5]Wilson (1968) also argues that group, i.e., syndicate, structures are important in finance and insurance problems.

random endowment of wealth), and information conditions that are identical to those in Townsend's model. However, Border and Sobel note (p. 533) that their arguments "use risk neutrality in an essential way" and that "it is not known if the monotonicity result ... extends to the risk averse case." This open question is particularly important for insurance applications of the model as risk aversion is typically thought to be the driving force behind most insurance arrangements. Further, we argue in Section 4 that the multilateral framework is important for many (costly) auditing problems.

The monotonicity result establishing the optimality of transfers and monitoring probabilities that are monotonically decreasing functions of agents' wealth reports is important because it is consistent with the following stylized facts. In insurance markets, a large loss can be viewed as a low wealth realization. Thus, a monotonic contract implies that policy holders receive higher transfers when they claim larger losses, and the probability of being audited is correspondingly higher for such reports. Border and Sobel (1987, p. 531) note that the tax interpretation of the monotonicity result is "more subtle." Taxes can be viewed as negative transfers and low wealth reports can be viewed as high itemized deduction claims. Thus, the monotonicity result implies that larger (total) tax payments are associated with larger wealth reports, but the probability of a tax audit is decreasing in reported wealth. The key insight is that a low wealth *claim,* not low wealth itself, makes a tax audit more likely.

The purpose of this paper is to generalize the costly state verification model to resolve the difficulties noted by Townsend and Border and Sobel that preclude its application to economic problems involving risk aversion and/or multiple agents (e.g., financial intermediation, insurance, and tax problems). Thus, we specify a model with multiple asymmetrically informed agents who have access to a costly state verification technology. We study the nature of contracts that support (information and resource) constrained Pareto efficient consumption allocations in the model when contracts are not restricted a priori to be symmetric, bilateral, or separable in endowments, and agents may be risk averse. We show that even in this more general environment, "debt-like" securities remain optimal when the monitoring technology is deterministic (i.e., the optimal multilateral contract has a lower interval), and transfer functions and monitoring probabilities remain monotonically decreasing functions of wealth when the monitoring technology is stochastic.

We use abstract measure theoretic arguments to derive the form of the optimal contracts in our more general economic setting. The key problem when agents are risk averse and contracts are explicitly multilateral is that non-trivial interdependencies among agents exist. Thus, strong measure theoretic tools such as the Isomorphism Theorem (cf., Section 3) and Lusin's Theorem (cf., Section 4) appear to be necessary to solve this interdependency problem. These tools allow us to change contracts in such a way that only the expected utility of one agent is affected while the expected utility of all other agents remains the same. We are then able to show that given any arbitrary initial contract, unless we start with lower intervals as monitoring sets (when monitoring is deterministic), or monotonically decreasing transfer and monitoring functions (when monitoring is stochastic), at least one agent can be made better off, which contradicts the (constrained) Pareto optimality of the arbitrary initial contract. These results are stated formally in Theorem 1, and Theorem 2 and Corollary 1, respectively.

2 The Model

Consider a two period exchange economy with finitely many individuals indexed by $i = 1, \ldots, n$. Each trader is described by a von Neumann-Morgenstern utility function, u_i, which is defined over second period consumption, c_i, and a random endowment, X_i. Let u_i be concave and monotonically increasing in consumption. Assume that the X_i are independent random variables. Denote a particular realization of X_i by x_i, let F_i be the distribution of X_i, let F^n be the joint distribution of $X_1, \ldots, X_n$, and assume that all distributions are non-atomic.[6] Finally, to ensure non-negative consumption, assume that the support of F_i is contained in $[m, \infty)$, where $m > 0$. The information conditions are as follows (cf., Townsend (1979, p. 281)). Each agent i privately observes the realization of his/her endowment, X_i, ex-post, but all agents have access to a costly state verification technology that can be used to *publicly* announce the realization to all other agents.

Let $\phi_i(\cdot)$ be the cost incurred by agent i from using the verification technology.[7] Denote by $t_i(x_1, \ldots, x_n)$ the net transfer function of agent i,

[6] A distribution is non-atomic if every single point has probability zero. This follows automatically if the distribution has a density.

[7] Verification is perfect in the sense that after monitoring occurs the true endowment

which describes the payment between the coalition and each agent i. This payment may be positive (indicating a state-contingent payment from the coalition to the agent), negative (indicating a payment by the agent to the coalition), or zero. Throughout our analysis we assume that agents' verification costs are an arbitrary positive function of the transfer payments, $\phi_i(t_i(\cdot))$. Because transfers need not be identical across agents, verification costs may differ as well. Townsend (1979, p. 269) considers two verification cost specifications, and our cost function includes both as special cases. In his first case, the verification cost is a fixed constant, and hence independent of the actual realization. In the second case, the verification cost of agent i depends on the transfer t_i, where the costs are strictly monotonic.

Resources are allocated in this economy via binding contracts. At time zero, agents have the opportunity to write contracts to provide for consumption next period. The structure of optimal contracts that emerge depends on the specification of agents' preferences, the distributions of random variables, the verification technology, and the nature of information in the economy. Three alternative ex-post information conditions are possible:

- When $\phi_i(t_i(\cdot)) \equiv 0$ for $i = 1, \ldots, n$, there is costless, and consequently complete public information about the realization of each X_i ex post. When agents have identical utility functions and weights and if the X_i are identically distributed, it follows that

$$c_i(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

  for all $i = 1, \ldots, n$. See Caspi (1978, p. 270, Theorem 2) for a formal proof of this result for the core.
- When $\phi_i(t_i(\cdot)) \equiv \infty$ for $i = 1, \ldots, n$, information is infinitely costly, so no verification is undertaken and information about each X_i remains completely private. The optimal multilateral contract which is individually rational in this case is autarky.
- In the remainder of the paper, we characterize the nature of optimal multilateral contracts under deterministic and stochastic verification, respectively, when information need not be entirely private nor public.[8]

---

is publicly reported without error.

[8]Existence of optimal contracts for the model with either deterministic or stochastic monitoring follows from Krasa and Villamil (1991b).

3 The Case of Deterministic Verification

In this section we study the form of Pareto efficient multilateral contracts that arise among agents under deterministic monitoring. Note that transfers, $t_i(\cdot)$, can be contingent only on endowment realizations of agent i which are publicly verified. In private information states, all transfers must be non-contingent. Let S_i denote the set of all announced realizations of X_i for which verification occurs, and let S_i^c denote the complement of S_i. We begin by defining a multilateral contract for this economy.

Definition 1. *A multilateral contract with deterministic verification for each agent $i = 1, \ldots, n$ is a pair (t_i, S_i), where $t_i(x_1, \ldots, x_n)$ is a net-transfer function for agent i from $I\!R^n$ into $I\!R$ and S_i is a set of endowment realizations announced by agent i for which monitoring occurs (with probability one). If agent i is verified, the endowment becomes public information.*

We restrict the analysis to the class of incentive compatible contracts:

Definition 2. *A collection of multilateral contracts (t_i, S_i) with deterministic verification is incentive compatible if $S_i = \bar{S}_i$ and $t_i(\cdot) = \bar{t}_i(\cdot)$ for every $i = 1, \ldots, n$, where (t_i, S_i) denotes the pre-state contractual commitment and $(\bar{t}_i, \bar{S}_i)$ denotes the post-state outcome.*

Definition 2 indicates that under an incentive compatible contract, agents do not misrepresent their private information (i.e., pre-state commitments are fulfilled ex post). Townsend (1988, pp. 416–418) uses a revelation principle argument to prove that incentive compatibility can be imposed without loss of generality. The following conditions generalize the incentive compatibility specification of Lemma 5.1 in Townsend (1979):

(IC1) $x_i \mapsto t_i(x_1, \ldots, x_i, \ldots, x_n)$ is constant on S_i^c, for a.e. x_j, $j \neq i$.

(IC2) $t_i(x_1, \ldots, x_i, \ldots, x_n) - \phi_i(t_i(\cdot))) \geq t_i(x_1, \ldots, y, \ldots, x_n))$; for a.e. $x_i \in S_i$, for every $y \in S_i^c$, and for a.e. x_j, $j \neq i$.

IC1 says that when agent i's endowment announcement is not verified (ceteris paribus), his/her net-transfer is constant (because transfers cannot depend on private information). IC2 says that it is (at least weakly) optimal for agent i to request verification when the endowment realization is in the verification set. Thus, it ensures that agent i requests verification when $x_i \in$

S_i. We assume that the incentive constraints are satisfied a.e. Thus, there exists a set of realizations of the agent's endowment which has measure zero in which it might be optimal to misreport. See Section 5 for a discussion of implementation, alternative specifications, and institutional interpretations of the incentive constraints.

We now state an information constrained optimization problem whose solutions characterize optimal multilateral contracts. The objective is to choose Pareto efficient net transfer functions, $t_i(X_1, \ldots, X_n)$, and sets of endowment realizations for which verification occurs, S_i, to maximize a weighted average of agents' utilities, subject to feasibility and information constraints. The λ_i denote weights on agents' utility functions.

Problem 3.1. *Choose t_i and S_i for $i = 1, \ldots, n$ to maximize*

$$\sum_{i=1}^{n} \lambda_i \int u_i \left[c_i(x_1, \ldots, x_n)\right] dF^n(x_1, \ldots, x_n), \tag{3.1}$$

subject to

$$0 \leq c_i \leq x_i + t_i(x_1, \ldots, x_n) - \phi_i(t_i(\cdot)) \text{ a.e. for all } i, \tag{3.2}$$

$$\sum_{i=1}^{n} t_i \leq 0 \text{ a.e.}, \tag{3.3}$$

$$t_i \text{ is incentive compatible for every } i \tag{3.4}$$

$$S_i \text{ is a measurable set for every } i. \tag{3.5}$$

The optimal multilateral contract maximizes the expected utility of all agents (3.1), subject to: (3.2) a budget constraint for each agent which holds almost everywhere; (3.3) an aggregate feasibility constraint which holds almost everywhere; (3.4) incentive-compatibility conditions IC1 and IC2; and (3.5) a standard measurability condition.

The purpose of this section is to characterize the nature of optimal contracts when verification is deterministic. Our main result is that (constrained) Pareto efficient multilateral contracts have lower interval monitoring sets, except for nullsets. Thus, we show that there exists a γ_i such that $S_i = [m, \gamma_i)$ for all i, except for a set of measure zero, where the lower interval may be trivial. Because monitoring is deterministic, it follows immediately from this result that the transfer function is constant for all $x_i \in S_i^c$ (for fixed x_j, $j \neq i$). As we noted at the outset, Townsend (1979, p. 283) proves a related lower interval result under several exogenous restrictions which he describes as "unpleasant" because they are necessary for technical reasons, but are not motivated by economic considerations. Specifically, he assumes:

(i) all transfers and verification costs are symmetric;

(ii) all trades are bilateral; and further

(iii) when both agents are verified, the transfer function is separable in endowment realizations (i.e., in our notation $t(x_1, x_2) = \hat{t}_1(x_1) + \hat{t}_2(x_2)$).[9]

Before beginning our formal analysis we describe the relationship between our result and Townsend's, and give an overview of the proof of Theorem 1.

Townsend specifies an optimization problem which involves the maximization of a weighted average of utilities, subject to information and resource constraints. However, instead of characterizing t_i and S_i directly as in our Problem 3.1, Townsend reformulates an analog of Problem 3.1 as a standard constrained maximization problem. The key difference between our approaches is that the maximizer in his reformulated problem is a function of only one variable. This follows from restrictions (i) and (iii), as they immediately imply that the transfer function is of the form $t(x_1, x_2) = \hat{t}(x_1) + \hat{t}(x_2)$. Under these restrictions it is only necessary to choose a one-dimensional transfer function, $\hat{t}$. Townsend considers the multilateral case (pp. 278–283) but reduces it to a similar one-dimensional problem by using (ii). This approach has two limitations. First, it precludes certain types of agent heterogeneity (i.e., (i) rules out transfer and cost function differences). Second, even when agents' transfer and cost functions are identical, restrictions (ii) and (iii) preclude certain economically plausible risk-sharing arrangements as noted in the Introduction.

In contrast, we characterize the solutions to Problem 3.1 directly. The maximizers are explicitly multi-dimensional transfer functions and verification sets, where transfer and verification cost functions need not be symmetric. We use abstract measure theoretic arguments to obtain our results, and these mathematical tools appear to be essential in our more general setting. We proceed as follows: Our main result in this Section is Theorem 1, which establishes that in a multi-agent economy with deterministic costly state verification, *all* solutions to Problem 3.1 have lower interval verification

[9]Separability is equivalent to the slope of the net-transfer function of agent i depending only on agent i's realization. This precludes most interesting externalities among agents.

sets (except for sets of measure zero). We prove the Theorem indirectly by assuming that there exists some arbitrary initial contract ($t_i(\cdot)$, S_i) which is optimal but is not a lower interval. We then define a measure preserving mapping, g, which allows us to transform the transfer functions, monitoring sets, and monitoring costs associated with the initial contract into an alternative contract (t'_i, S'_i) such that the new contracts are feasible, incentive compatible, strictly increase the expected utility of at least one agent, and leave the expected utility of all other agents unaffected. This contradicts the optimality of the original (non-lower monitoring interval) contract, hence it establishes the optimality of contracts with lower monitoring intervals.

Roughly speaking, we contradict the optimality of non-lower intervals in the following way. We move a part of the original (non-lower interval) monitoring set of one agent (say agent one) to the left, mapping it into a set where there was previously no state verification. Such sets (with positive measure) always exist if the initial contract was not a lower monitoring interval, and we construct these sets to be compact. The existence of a measure preserving one-to-one mapping, g, between these two sets follows from the Isomorphism Theorem which says that measure preserving one-to-one mappings exist between all separable and complete measure spaces (where both spaces have the same measure). Since compact subsets of $\mathbb{R}$ are separable and complete (in the induced topology) the Theorem can be applied.

Feasibility and incentive compatibility of the alternative contract are straightforward to show because g is measure preserving and one-to-one. It is also reasonably straightforward to show that the expected utility of agent one increases by a Rothschild and Stiglitz increasing risk argument. Townsend (1979, p. 288) uses a similar argument in the proof of Proposition 3.2, which is his lower-interval result for two-agents, one risk neutral, with *fixed* monitoring costs. Thus, the reader may wonder why we use abstract measure theory to obtain our results. The remaining and key step in the proof is to show that the utility of all other agents does not decrease under the alternative contract. In Townsend's proof, this follows immediately from risk neutrality and fixed verification costs.[10] In our setting with multiple risk-averse agents

[10]Townsend (1979, p. 287, Proposition 3.1) proves a second lower interval result for a *bilateral* contracting problem where agents may be risk averse and the monitoring cost function is convex with $\phi_i(0) < 1$. The Euler equation argument he uses to obtain his result depends crucially on restrictions (i), (ii), and (iii). It does not appear that this approach can be readily extended to the multilateral case because of the interdependency

and arbitrary verification cost functions his argument breaks down exactly at this step because all contracts are *interdependent.* Without an additional argument, it is not possible to avoid affecting other agents' expected utility nor to see in which direction their utilities change. Measure preserving mappings impose the necessary structure to overcome this problem.

We begin our analysis by defining a measure preserving mapping. As indicated above, this concept is crucial for the arguments that follow.

Definition 3. *Let (Y_i, β_i, μ_i), $i = 1, 2$ be two measure spaces and let $g: Y_1 \to Y_2$ be a measurable function. For every $A \in \beta_2$ define $gA = \{ga: a \in A\}$. Then g is measure preserving iff $\mu_1(g^{-1}A) = \mu_2(A)$.*

The following Remark is an immediate consequence of Definition 3.[11]

Remark 1. *Let f be an integrable function on Y_2, and let g be a measure preserving transformation as defined above. Then $f \circ g$[12] is integrable and the following holds:*

$$\int_{Y_2} f(x)\, d\mu_2(x) = \int_{Y_1} f(g(x))\, d\mu_1(x).$$

Remark 1 corresponds to Theorem 1.6.12 of Ash (1972) or Remark 28.14 of Parthasarathy (1977). For completeness we give the proof in the Appendix. This Remark is essential for the proofs of our main results as it establishes that whenever we change the payoffs to one agent in a measure preserving way (i.e., choose a measure preserving function g), then the expected utility from an arbitrary initial contract $t_i(X_1, \ldots, X_n)$ and a transformed alternative contract $t_i(g(X_1), X_2, \ldots, X_n)$ is the same for all other agents.

---

problem.

[11]Consider the following example of a measure preserving mapping. Let $Y_1 = [0,1] \cup \{2\}$ and $Y_2 = [1, 2]$. On both sets consider the standard Lebesgue measure. Then the function

$$g(x) = \begin{cases} x+1 & \text{if } x \in [0,1]; \\ 1 & \text{if } x = 2; \end{cases}$$

is measure preserving in this example (though not a one-to-one mapping).

[12]$f \circ g$ is the composition of f and g, i.e. $f \circ g(x) = f(g(x))$.

To construct measure preserving mappings we use the Isomorphism Theorem from measure theory (cf., Parthasarathy (1977) Proposition 26.6).

Isomorphism Theorem. *Let Y_i, $i = 1, 2$, be complete and separable metric spaces, and let μ_i be non-atomic Borel measures on Y_i such that $\mu(Y_1) = \mu(Y_2) > 0$. Then the two measure spaces are isomorphic, i.e., there exist two sets of measure zero, N_i, $i = 1, 2$, and there exists a measure preserving transformation, $g: Y_1 \setminus N_1 \mapsto Y_2 \setminus N_2$, whose inverse exists and is also measure preserving.*[13]

We now state our main result concerning the nature of optimal contracts in a multi-agent economy with deterministic costly state verification.

Theorem 1. *Assume that the utility functions of all agents are twice continuously differentiable and that $u'' < 0$. Furthermore assume that $\phi_i(t) > 0$ for every agent i and for every $t \in \mathbb{R}$. Let the endowments of the agents be described by independent random variables X_i for all $i = 1, \dots, n$. Then all solutions to Problem 3.1 have lower interval verification sets, except for sets of measure zero (i.e., there exists a γ_i such that $S_i \triangle \{X_i: X_i < \gamma_i\}$ has measure zero.)*[14]

Proof. We proceed indirectly. Without loss of generality, assume that the monitoring set of agent one is not a lower interval. Let μ be the distribution of the endowment of agent one. Then there exist compact sets K_i, $i = 1, 2$ with positive measure, and such that $k_1 < k_2$ for all $k_i \in K_i$ and such that $K_1 \subset \mathbb{R} \setminus S_1$ and $K_2 \subset S_1$. By regularity[15] and non-atomicity of the measure, we can assume that $\mu(K_1) = \mu(K_2)$. Note that the K_i are separable and complete because they are compact. Thus, by the Isomorphism Theorem there exists a measure preserving mapping $h: K_1 \setminus N_1 \to K_2 \setminus N_2$ such that h^{-1} exists and is also measure preserving, where N_i, $i = 1, 2$ are sets of measure

[13] "$\setminus$" denotes set theoretic subtraction.

[14] "$\triangle$" denotes the symmetric difference: $A \triangle B = (A \setminus B) \cup (B \setminus A)$, for arbitrary sets A and B.

[15] Regularity means that $\mu(A) = \inf\{\mu(O): O \supset A,\ O \text{ open}\} = \sup\{\mu(F): F \subset A,\ F \text{ closed}\}$. Our measure μ is regular, since every probability measure on a metric space is regular (cf., Parthasarathy (1977) Proposition 19.13).

zero. Note that h can be extended to $I\!R$ by

$$g(x) = \begin{cases} h(x) & \text{if } x \in K_1 \setminus N_1; \\ h^{-1}(x) & \text{if } x \in K_2 \setminus N_2; \\ x & \text{otherwise.} \end{cases}$$

Clearly, g is again measure preserving.

Recall that $t_i(x_1, \ldots, x_n)$ are transfer functions associated with some arbitrary initial contract, where the monitoring set of agent one is not a lower interval. Thus for every agent i, now define new transfers t_i' by

$$t_i'(x_1, \ldots, x_n) = t_i(g(x_1), x_2, \ldots, x_n).$$

Further, define the new monitoring set of agent one by $S_1' = g^{-1}(S_1)$ and $S_i' = S_i$ for $i = 2, \ldots, n$. The strategy of the proof is to show the following: (i) The transfer functions associated with the new contracts $(t_i'(\cdot), S_i')$ are feasible; (ii) the new contracts are incentive compatible; (iii) the utility of all other agents $i \neq 1$ does not change; and (iv) the utility of agent one strictly increases. This gives the contradiction to the assumed optimality of a non-lower interval contract. (i)-(iv) are proved as follows:

(i) Let $A = \{(x_1, \ldots, x_n) \colon \sum_{i=1}^n t_i(x_1, \ldots, x_n) > 0\}$. Define $\tilde{g}$ on $I\!R^n$ by $(x_1, \ldots, x_n) \mapsto (g(x_1), x_2, \ldots, x_n)$. Clearly, $\tilde{g}$ is measure preserving with respect to the joint distribution of the X_i. Then, $\tilde{g}^{-1}A = \{(y_1, \ldots, y_n) \colon g(y_1) = x_1;\ y_i = x_i$ for all $i > 1$, and $\sum_{i=1}^n t_i(x_1, \ldots, x_n) > 0\} = \{(y_1, \ldots, y_n)\colon$ such that $\sum_{i=1}^n t_i(g(y_1), y_2, \ldots, y_n) > 0\}$. Since $\tilde{g}$ is measure preserving, (3.3) implies that $\tilde{g}^{-1}A$ has measure zero. Hence,

$$\sum_{i=1}^n t_i(g(x_1), x_2, \ldots, x_n) \leq 0 \text{ a.e.}$$

which proves feasibility.

(ii) Incentive compatibility requires IC1 and IC2 to be fulfilled. IC1 is obvious. Let $\bar{t}_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ denote the constant payment to agent i in non-monitoring states. We first show that IC2 is satisfied for $i \geq 2$, (the argument is similar to that given for (i)). Define $\tilde{g}$ as above, but now let $A = \{(x_1, \ldots, x_n) \colon t_i(x_1, \ldots, x_n) - \phi_i(t(x_1, \ldots, x_n) < \bar{t}_i(x_1, \ldots, x_n)\}$. Then it follows that $\tilde{g}^{-1}A = \{(x_1, \ldots, x_n) \colon t_i(g(x_1), \ldots, x_n) - \phi_i(t(g(x_1), \ldots, x_n)) < \bar{t}_i(g(x_1), \ldots, x_n)\}$. Since $\tilde{g}$ is measure preserving, IC2 implies that $\tilde{g}^{-1}A$ has measure zero. Hence IC2 holds for the new contract for all agents $i \geq 2$. It remains to give the proof for $i = 1$. This, however, follows immediately from the argument for $i \geq 2$ and the fact that

$$\begin{aligned}\bar{t}'_1(x_2, \ldots, x_n) &= \sup_{y_1 \in S_1'^{\,c}} t_1(g(y_1), \ldots, x_n) \\ &= \sup_{y_1 \in S_1^c} t_1(y_1, \ldots, x_n) = \bar{t}_1(x_2, \ldots, x_n),\end{aligned}$$

because g is one-to-one. This proves (ii).

(iii) Apply Remark 1 and Fubini's Theorem (cf., Ash (1972), Theorem 2.6.4). Let c_i' denote consumption under the new contract, and let c_i denote consumption under the original contract for agent i. Note that for every $i \neq 1$ we have $c_i(g(x_1), x_2, \ldots, x_n) = c_i'(x_1, \ldots, x_n)$. We must show that $\int u_i(c_i)\, dF^n = \int u_i(c_i')\, dF^n$, which means that the expected utilities are the same. This follows from Fubini's Theorem since

$$\begin{aligned}&\int\!\!\int \cdots \int u_i(c_i(g(x_1), x_2, \ldots, x_n))\, dF_1(x_1) dF_2(x_2), \ldots dF_n(x_n) \\ &\qquad = \int\!\!\int \cdots \int u_i(c_i(x_1, x_2, \ldots, x_n))\, dF_1(x_1) dF_2(x_2) \ldots dF_n(x_n).\end{aligned}$$

Equality follows from Remark 1, i.e., the fact that g is measure preserving. This proves (iii).

(iv) For given $(x_2, \ldots, x_n)$ define

$$f(x_1) = t_1(x_1, \ldots, x_n) - \phi_1(t_1(x_1, \ldots, x_n)).$$

Because of IC1 and IC2, transfers (net of monitoring costs) in monitoring states are always higher than transfers in non-monitoring states. g moves these high transfers to the left (i.e., to low income states) and vice versa.[16] By Lemma 2 in the Appendix, agent one is strictly better off under the new contract. This contradicts the assumed optimality of the original contract, proving the Theorem.

4 The Case of Stochastic Verification

In this section we study the form of Pareto efficient multilateral contracts that arise among agents under stochastic monitoring. We begin by defining a multilateral contract for this economy.

[16]That is $f(k_1) < f(k_2)$ for every $k_1 \in K_1$ and for every $k_2 \in K_2$. This is exactly the condition under which Lemma 2 holds.

Definition 4. *A multilateral contract with stochastic verification for each agent $i = 1, \ldots, n$ is a pair (t_i, p_i), where $t_i(x_1, \ldots, x_n)$ is a net-transfer function for agent i from $\mathbb{R}^n$ into $\mathbb{R}$, and $p_i: [m, \infty]^n \to [0,1]$ is a function which indicates the probability that agent i's endowment announcement is verified. If agent i is verified, the endowment becomes public information.*

Alternative formulations of the stochastic monitoring problem have been studied previously by other authors. In particular, Townsend (1988, p. 424) reports the results of systematic numerical analyses of costly state verification economies with stochastic monitoring and gives examples of non-monotonic monitoring probabilities. His results stem from the fact that the monitoring probability function, p_i, in his model is defined on $[m, \infty]$. That is, whether or not an agent is verified depends only on the agent's own announcement, and is independent of all other agents' announcements.[17] In contrast, in our model monitoring depends on the agent's own endowment announcement *and* on the announcements of all other agents (i.e., p_i is defined on $[m, \infty]^n$ in Definition 4). This specification seems reasonable for the stochastic auditing applications of the model described at the outset. For example, the probability of a tax audit by the IRS is related not only to an individual's own income tax return, but also to the returns filed by all other individuals in the economy.[18] Finally, Border and Sobel also prove a monotonicity result. However, as we noted in the Introduction, their arguments depend crucially on risk neutrality.

Our main goal in this section is to characterize the solutions to an information constrained optimization problem with stochastic monitoring. Before beginning our formal analysis we first discuss an inherent difficulty that emerges in economies with stochastic monitoring *and* risk averse agents.[19] The problem stems from the fact that stochastic monitoring generates addi-

[17]Townsend's example is for a discrete (hence atomic) distribution. However, because it is an equal distribution our proof immediately goes through (but it breaks down for discrete distributions which are not equal distributions). We are not aware of an example where the "discreteness" is solely responsible for the non-monotonicity.

[18]That is, an individual with a university salary is more likely to be audited in a small college town (Urbana, IL) than in the Silicon Valley (Palo Alto, CA).

[19]Randomness is inherent in a stochastic monitoring technology. Thus, one may interpret the optimal consumption allocations derived from Problem 4.1 as "consumption lotteries," but they differ from the consumption lotteries in Prescott and Townsend (1984) (which are introduced as a device to obtain a concave programing problem).

tional uncertainty into expected consumption allocations, and this additional uncertainty decreases the expected utility of risk averse agents.[20] The key problem is that states with low endowment realizations are the same states where the probability of monitoring is the highest. These high variance states are precisely the states of most concern to risk averse agents. In general it is difficult to precisely characterize the marginal loss of utility to an agent from the additional uncertainty caused by stochastic monitoring. Transfers which are contingent not only on all agents' endowment realizations (as they are in our model), but also on whether or not monitoring is actually performed (which does not occur in our model) might ameliorate the negative utility effects associated with stochastic monitoring somewhat. However, it is unlikely that such transfers would eliminate these effects entirely.

We consider two polar cases which are designed to address the "marginal utility loss" problem experienced by risk averse agents. We first consider the case where monitoring costs are borne by each individual agent, but restrict agents' utility functions to be separable in consumption and monitoring cost. This approach is often employed in the literature (e.g., Moohkerjee and Png (1989)), hence we use it in the statement of Problem 4.1 below. However, our proofs also apply to an alternative specification where agents are able to diversify their individual specific monitoring cost risk (e.g., if monitoring occurs, the monitoring costs of agent i are borne by all other agents $i \neq j$). We defer discussion of this second specification until after we have proved our main results (Theorem 2 and Corollary 1).

We now state the optimization problem for this economy:

Problem 4.1. *Choose $t_i(\cdot)$ and $p_i(\cdot)$ for $i = 1, \ldots, n$ to maximize:*

$$\sum_{i=1}^{n} \lambda_i \int \left[v_i\left(x_i + t_i(\cdot)\right) - p_i(\cdot)\phi_i(t_i) \right] dF^n(x_1, \ldots, x_n), \tag{4.1}$$

subject to

$$0 \leq c_i \leq x_i + t_i(x_1, \ldots, x_n) \text{ a.e. for all } i, \tag{4.2}$$

$$\sum_{i=1}^{n} t_i \leq 0, \text{ a.e.} \tag{4.3}$$

[20]Note that stochastic monitoring also has the countervailing beneficial effect of reducing expected monitoring costs (relative to deterministic monitoring).

$$
\begin{aligned}
&v_i(x_i + t_i(x_1,\ldots,x_i,\ldots,x_n)) - p_i(x_1,\ldots,x_n)\phi_i(t_i(\cdot)) \geq \\
&\qquad (1 - p_i(x_1,\ldots,y,\ldots,x_n))v_i(x_i + t(x_1,\ldots,y,\ldots,x_n)) \\
&\qquad\qquad + p_i(x_1,\ldots,y,\ldots,x_n)[v_i(0) - \phi_i(t_i)],\ \textit{for all } i, \\
&\textit{for all } y,\ \textit{and for a.e. } x_i;\ \textit{and} \qquad (4.4) \\
&0 \leq p_i(x_1,\ldots,x_n) \leq 1,\ \textit{for every } x_i. \qquad (4.5)
\end{aligned}
$$

Equation (4.1) reflects the consumption and monitoring cost separability restriction described previously. Separability implies that each agent's utility from consumption is independent of the non-pecuniary (effort cost) imposed on the agent by the monitoring procedure. Loosely speaking, the idea is that the monitoring process causes no additional utility or disutility other than the direct costs. Equation (4.3) is the same as in Problem 3.1. Equation (4.4) is the incentive compatibility constraint under stochastic monitoring.[21] The left-hand side of (4.4) is the expected utility of agent i from truthfully reporting endowment realization x_i; and the right-hand side is the expected utility of agent i from announcing any other realization $y \neq x_i$. When agent i misreports and is verified, he/she receives a zero transfer and the entire endowment is confiscated, so utility is $v_i(0) - \phi_i(t_i(\cdot))$. We have implicitly assumed that it is optimal to punish an agent as much as possible (by seizing the entire endowment) for misreporting. This, however, is straightforward to show since maximizing the penalty minimizes the propensity to cheat. We again refer the reader to Section 5 for further discussion of incentive compatibility. Finally, (4.5) states that the p_i are probabilities.

We now give an overview of the proof of Theorem 2. This Theorem shows that the transfer function associated with the optimal contract is a decreasing function of wealth when monitoring is stochastic. As in Theorem 1, we proceed indirectly: Assume that the transfer function of an agent (say agent one) is not a monotonically decreasing function of wealth over the entire support of the distribution. We again wish to use the Isomorphism Theorem to find a measure preserving one-to-one function g which maps arbitrary initial contracts into an alternative contract which is "more monotonic."[22] We show that this "more monotonic" alternative contract: (i) is feasible;

[21]Townsend (1988, pp. 416–418) uses a revelation principle argument to prove that this restriction can be imposed without loss of generality.

[22]In general it is not possible (even for very simple cases) to construct a monotonic contract directly with a measure preserving transformation. Consider the following example: Choose the interval $[0,1]$ with the standard Lebesgue measure. Let $f(x) = x(1-x)$. Now

(ii) is incentive compatible; (iii) does not decrease the expected utility of all other agents; and (iv) strictly increases the expected utility of agent one. This establishes the optimality of contracts with monotonically decreasing transfer functions.

The first step of the proof, since the argument is indirect, is to establish a uniform violation of (decreasing) monotonicity of an arbitrary initial (non-monotonic) transfer function. We begin by showing that it is possible to find two compact sets with positive measure, denoted $\mathcal{U}$ and $\mathcal{V}$, where $\mathcal{U}$ is strictly to the left of $\mathcal{V}$, and such that all values of the transfer function in $\mathcal{U}$ are strictly below the values which the transfer function assumes in $\mathcal{V}$.[23] To construct such sets, we use Lusin's Theorem (cf., Parthasarathy (1977) Proposition 24.21 and Corollary 24.22), which says that for any integrable function (on a complete and separable metric space) there exist arbitrary large compact subsets of the domain such that the restriction of a function on this compact subset is continuous. We use this continuity to establish the desired (uniform) violation of monotonicity of the transfer function on $\mathcal{U}$ and $\mathcal{V}$. The main insight in this part of the proof is that it is not sufficient to establish a violation of monotonicity of the transfer function for single points as the analysis necessarily excludes sets of measure zero. Hence, starting with two points z_1, z_2 for which monotonicity is violated, we must establish a violation which also holds on a set of positive measure contained in neighborhoods of these two points. For continuous functions this is obviously always the case. Fortunately, Lusin's Theorem implies that this is also true almost everywhere for arbitrary measurable functions (by continuity of such a function on compact subsets).

The remainder of the proof is similar in structure to Theorem 1: We apply a version of the Isomorphism Theorem (proved in Lemma 3) to get a

assume (indirectly) that there exists a measure preserving transformation g on $[0,1]$ such that $f \circ g(x) = g(x)(1 - g(x))$ is monotonic. The function is quadratic, so there are two solutions x_i, $i = 1, 2$ to any equation $x(1-x) = z$. Hence, there exist $x_1 \neq x_2$ such that $f \circ g(x_1) = f \circ g(x_2)$. Assume that $x_1 < x_2$. Since $f \circ g$ is monotonic, f is constant on the image of the interval $[x_1, x_2]$ under g. This, however, means that $g([x_1, x_2])$ contains at most two points. This is a contradiction to g being measure preserving.

[23]For technical reasons we prove an even stronger violation of monotonicity. In particular, we show that if the endowment realization x lies in $\mathcal{U}$ and a transfer corresponding to an arbitrary state in $\mathcal{V}$ is used instead of the transfer $t_1(x)$, then the consumption of agent one strictly increases. A similar condition holds if the realized state is an element of $\mathcal{V}$.

measure preserving mapping h between the arbitrary initial (non-monotonic) contract and a (more monotonic) alternative contract, on the two compact sets $\mathcal{U}$ and $\mathcal{V}$. We then show that (i)–(iv) hold. However, unlike in Theorem 1 with deterministic monitoring, when we apply the Isomorphism Theorem in the stochastic case, we must apply it "slice-wise."[24] The basic problem is that the sets $\mathcal{U}$ and $\mathcal{V}$ do not necessarily have a product structure, i.e., we cannot represent $\mathcal{U}$ or $\mathcal{V}$ in the form $A \times C$ where $A \subset I\!R$ and $C \subset I\!R^{n-1}$ and $\mathcal{V}$ as $B \times C$ where $B \subset I\!R$. If the sets had the product structure, then we could apply the Isomorphism Theorem to construct a measure preserving mapping between A and B for agent one when the realizations of all other agents are fixed. In Lemma 3 we generalize the Isomorphism Theorem so that for fixed realizations $(x_2, \ldots, x_n)$ of all other agents we can still establish an isomorphism between respective "slices" of the sets A and B. We define the mapping on every slice in Lemma 3 in a way which ensures that we get a measure preserving "slice-wise" mapping, and then use Fubini's Theorem to get a measure preserving mapping h between the sets A and B. The technical problem in applying the argument is to ensure measurability of h, but this follows from a measurable selection Theorem (also contained in Lemma 3). The strategies of the arguments for (i), (ii), (iii), and (iv) remain similar to those used in Theorem 1.

We now state our main result concerning the nature of optimal contracts in a multi-agent economy with stochastic verification.

Theorem 2. *Let (t_i, p_i) for $i = 1, \ldots, n$ be a collection of Pareto optimal contracts. Then there exists a set of measure zero N such that for every agent i and for every $z_1 = (x_1, \ldots, x_i, \ldots, x_n)$, and $z_2 = (x_1, \ldots, y_i, \ldots, x_n)$ with $z_1, z_2 \in I\!R^n \setminus N$ it follows that $t_i(z_1) \geq t_i(z_2)$ if $x_i \leq y_i$, i.e., the transfers are monotonically decreasing a.e.*

Proof. We proceed indirectly. Without loss of generality we can assume that the transfer function of agent one is not monotonic a.e. Let $\mathcal{O}$ be the union of all open sets with measure zero. Then $\mathcal{O}$ itself is open and has measure zero. By Lusin's Theorem (cf. Ash (1972), Corollary 4.3.17(b)) there exists for

[24] A "slice" of an arbitrary set $A \subset I\!R$ is given by

$$A_{(x_2, \ldots, x_n)} = \{(x_1, x_2, \ldots, x_n) : (x_1, x_2, \ldots, x_n) \in A\}.$$

every $\varepsilon > 0$, a compact subset $K \subset I\!R^n$ with $\mu(I\!R \setminus K) < \varepsilon$ and such that t_1 is continuous on K. Without loss of generality we can assume that $\mathcal{O} \cap K = \emptyset$ (otherwise take $K \setminus \mathcal{O}$). Hence, we can construct an increasing sequence of compact sets K_i such that t_1 is continuous on each of the K_i and such that $I\!R^n \setminus \bigcup_{i=1}^{\infty} K_i$ has measure zero. Since t_1 is not monotonic a.e. there must exist $z^1 = (x_1, x_2, \ldots, x_n)$, and $z^2 = (y_1, x_2, \ldots, x_n)$ such that $x_1 < y_1$, and $t_1(z^1) < t_1(z^2)$, and such that $z^1, z^2 \in \bigcup_{n=1}^{\infty} K_n$. For a sufficiently large n we can assure that $z^1, z^2 \in K_n$. Thus, t_1 is continuous on K_n.[25] Choose γ_1 and γ_2 such that $x_1 + t_1(z^1) < \gamma_1 < x_1 + t_1(z^2))$ and $y_1 + t_1(z^1) < \gamma_2 < y_1 + t_1(z^2)$. Then there exist compact neighborhoods $\mathcal{U}$ of z^1 and $\mathcal{V}$ of z^2 such that

(a) $u_1 + t_1(u) < \gamma_1 < u_1 + t_1(v)$; and

(b) $v_1 + t_1(u) < \gamma_2 < v_1 + t_1(v)$,

for every $u \in \mathcal{U}$ and $v \in \mathcal{V}$, where u_1 and v_1 are the first coordinates of u and v, respectively. Furthermore, we can assume that $\mathcal{U}$ is to the left of $\mathcal{V}$, i.e., for every $u \in \mathcal{U}$ and for every $v \in \mathcal{V}$ we have $u_1 < v_1$. Since $\mathcal{U}$ and $\mathcal{V}$ are neighborhoods, they must have positive measure (since their intersection with $\mathcal{O}$ is empty). By the "generalized Isomorphism Theorem" (Lemma 3) there exist subsets $A \subset \mathcal{U}$ and $B \subset \mathcal{V}$ and measure preserving mappings $h^1: A \to B$ and $h^2: B \to A$ such that for fixed $(x_2, \ldots, x_n)$ the mappings $x \mapsto h^i(x_1, \ldots, x_n)$ are measure preserving on $A_{(x_2,\ldots,x_n)}$ and $B_{(x_2,\ldots,x_n)}$, respectively.[26]

Now define

$$f(x_1, \ldots, x_n) = \begin{cases} h^1(x_1, \ldots, x_n) & \text{if } x \in A; \\ h^2(x_1, \ldots, x_n) & \text{if } x \in B; \\ (x_1, x_2, \ldots, x_n) & \text{otherwise.} \end{cases}$$

Then for fixed $(x_2, \ldots, x_n)$ the mapping $x_1 \mapsto f(x_1, \ldots, x_n)$ is a measure preserving transformation on $I\!R^n_{(x_2,\ldots,x_n)}$, where $I\!R^n_{(x_2,\ldots,x_n)}$ is given by the set $\{(x_1, x_2, \ldots, x_n): x_1 \in I\!R\}$. Let g denote the first coordinate of $f(x) = (f_1(x), \ldots, f_n(x))$. Then $x_1 \mapsto g(x_1, \ldots, x_n)$ is a measure preserving transformation on $I\!R$ for fixed $(x_2, \ldots, x_n)$.

Now define new transfers denoted by $t_i(g(x_1, \ldots, x_n), x_2, \ldots, x_n)$ and new monitoring probabilities denoted by $p_i(g(x_1, \ldots, x_n), x_2, \ldots, x_n)$. We show

[25] Note that all neighborhoods are in the *induced* topology on K_n and not in the original topology of $I\!R^n$, i.e., $\mathcal{U}$ is a neighborhood of $x \in K_n$ if there exists a neighborhood $\mathcal{W}$ of x in $I\!R^n$ such that $\mathcal{U} = K_n \cap \mathcal{W}$.

[26] $A_{(x_2,\ldots,x_n)} = \{(x_1, x_2, \ldots, x_n): (x_1, \ldots, x_n) \in A\}$ and similar for B.

that these new contracts are: (i) feasible, (ii) incentive compatible, (iii) preserve the utility of all agents $i \neq 1$, (iv) increase the utility of agent one.

(i) Feasibility follows as in the proof of Theorem 1.

(ii) Incentive compatibility requires (4.4) to be satisfied. There are three possible cases. First, assume the true realization $(x_1, \ldots, x_n)$ lies in B. If it is profitable to cheat in this situation under the alternative contract, then it must also have been profitable with the initial contract in state $g^{-1}(x_1, x_2, \ldots, x_n)$, because the transfers are the same under the two contracts but under the initial contract the endowment of agent one was lower (hence the penalty if detected cheating was less severe). This contradicts incentive compatibility of the initial contract. Second, assume the realization lies in A. If it is profitable to cheat in this situation under the alternative contract, then it must have been even more profitable under the initial contract as the transfer was lower. This again contradicts optimality of the initial contract. Finally, for all other realizations the two contracts are the same. This proves (ii). However, also note that the monitoring probabilities can be reduced slightly without violating incentive compatibility.

(iii) The expected utility of all other agents is unchanged, cf. Theorem 1.[27]

(iv) Since (a) and (b) are fulfilled, and since A is to the left of B we can apply Lemma 1 for agent one for fixed $x_2, \ldots x_n$. Agent one is therefore strictly better off with the alternative contract (we exchange high transfers to low income states and vice versa). Since the monitoring probabilities can be reduced slightly without violating incentive compatibility because of (ii), the utility of agent one can be strictly increased. Thus, Fubini's Theorem implies that the expected utility of agent one is strictly greater under the alternative contract. Thus, contracts which are non-monotonic cannot be optimal. This proves the Theorem.

The following Corollary follows immediately from Theorem 2.

Corollary 1. *Under the assumptions of Theorem 2 it follows that $p_i(z^1) \geq p(z^2)$, i.e., the probabilities of verification are monotonically decreasing a.e. in endowments.*

[27]The utility of an agent depends on $x_1, \ldots, x_n$. Using Fubini's theorem, we can first integrate over the realization x_1 in order to compute the expected utility. However, for fixed $x_2, \ldots, x_n$, the mapping $x_1 \mapsto g(x_1, \ldots, x_n)$ is measure preserving. Thus, g drops out of the integral when integrating over x_1 (cf., Remark 1).

Proof. The Corollary follows immediately from the fact that the transfers are monotonically decreasing: Let $x_1 \leq y_1$. Consider two endowments $z^1 = (x_1, \ldots, x_n)$ and $z^2 = (y_1, x_2, \ldots, x_n)$, and assume that monotonicity of the probabilities is violated for agent one. By Theorem 2, $t_1(z^1) \geq t_1(z^2)$. Now choose the same probabilities of monitoring for z^1 and z^2, and suppose it were profitable for the agent to cheat in some other state and announce z^2. Then it would be at least as profitable to announce z^1 since the transfer is at least as high and the probability that cheating is detected is lower. This contradicts incentive compatibility of the original contract. Hence $p_1(z^1) \geq p_1(z^2)$.

We conclude this section by discussing the alternative monitoring cost specification described before the statement of Problem 4.1. That is, instead of assuming that each risk averse agent i privately bears the entire "utility loss" stemming from stochastic monitoring, Theorem 2 and Corollary 1 continue to hold if we assume that a mechanism exists whereby the monitoring costs of agent i are borne by all agents $j \neq i$ (when monitoring occurs). This follows from the fact that steps (i), (ii) and (iii) from the proof of Theorem 2 remain valid under either specification of the model because the transfers and monitoring probabilities have the same expected value *and* the same distribution (although we did not use this fact in the proof of Theorem 2 because of the assumed separability of the utility function). Examples of mechanisms in actual economies which appear to be qualitatively similar to this second (publicly borne) cost specification are tax surcharges (levied by a government) or a reduction in the "dividend credits" commonly rebated to policy holders by insurance companies (e.g., TIAA-CREF and many other insurance companies follow this practice).

5 Discussion of Results and Extensions

In this paper we generalize the costly state verification model to allow risk averse agents who need not be identical ex ante to write multilateral contracts. Bilateral versions of the model have proved useful for many economic applications, and we believe that this multilateral extension will expand the class of economic problems that can be addressed in this framework. Of course, whether a problem is best analyzed in a bilateral or multilateral contracting framework depends on its underlying economic structure. However, a multilateral version of the model seems necessary for many types of insurance problems and certain types of financial intermediation problems (cf., Boyd and Prescott (1986)). In the remainder of this section we will discuss the implications of our results and extensions.

We first focus on the incentive compatibility constraint used in our analysis. Townsend (1988) notes that in order to justify this restriction in costly state verification models it is necessary to formulate the underlying revelation game. This works as follows: Contracts are written before uncertainty is revealed. Uncertainty is then privately revealed, and each agent sends a message (i.e., reports a state). Thus, agents play a Nash game in messages where each agent has beliefs over whether all other agents tell the truth. When the analysis is restricted to truth-telling equilibria, it follows that each agent expects all other agents to tell the truth. In such a framework the pointwise incentive constraints commonly used in costly state verification models follow. This formulation implicitly contains a great deal of communication among agents, in the sense that decisions are made based on the expected announcements by all other agents.

The other extreme that we now consider is a game with no communication among agents. We are concerned with two issues. First, what are the implications of such an environment for the form of the incentive constraints. Second, which environment (one with communication or one with no communication) seems most plausible for the economic problems which motivate this paper. We begin with the first issue. In a game with no communication among agents, each agent makes an announcement with no knowledge of other agents' announcements. This corresponds to a Harsanyi (1967) type Bayesian Nash game, where the incentive constraints need not hold pointwise but only in expected value.[28] Theorem 2 and Corollary 1 immediately

[28]IC1 would be written: $x_i \mapsto \int t_i(x_1, \ldots, x_i, \ldots, x_n)\, dF(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is constant on S_i^c; and IC2 would be written:

$$\int t_i(x_1, \ldots, x_i, \ldots, x_n) - \phi(\cdot)\, dF(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$
$$\geq \int t_i(x, \ldots, y, \ldots, x_n)\, dF(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n);$$

for all $x_i \in S_i$ and for every $y \in S_i^c$. In both cases $dF(\cdot)$ denotes integration with respect to the joint distribution of the random variables X_j, $j \neq i$. Unlike in the pointwise specification, these constraints need only hold on average.

go through under this alternative formulation of the constraint because we do not use incentive compatibility in any essential way in the proof. Rather, we need only check that it remains satisfied.[29] From a technical point of view, this step of the proof requires us to show that our construction does not move us out of the set of all incentive compatible contracts—and this is of course easier to show if the constraint set is bigger. Thus, the expected value form of the incentive constraint does not change the structure of the optimal contract in an environment with stochastic monitoring. In fact, it facilitates the technical arguments necessary to prove the result.

In contrast, in Theorem 1 we again check that incentive compatibility conditions IC1 and IC2 are satisfied in step (ii) of the proof, but we also use these conditions in step (iv) in an essential way. In particular, we use them in (iv) to show that the transfer in every non-monitoring state is always higher than the transfer in every monitoring state. Thus, the final step in the proof of Theorem 1 does not go through with an incentive constraint which holds only in expected value. In fact, it turns out that under the mathematically weaker expected value constraint, the transfers associated with the optimal contract need no longer be constant on the non-monitoring set. We first show this in a simple (but not pathological) example and then provide an economic interpretation of the result.

Example 1. Consider a discrete distribution and two agents: agent one is risk neutral and agent two is very risk averse. The same kind of example also goes through for continuous distributions and if one agent is (slightly) risk averse. Assume that there are four states which occur with equal probability. The endowment of agent one is given by $(7,7,3,3)$ and of agent two by $(7,3,7,3)$. Clearly, the two endowments are independent. Let ϕ be a constant monitoring cost. Choose $S_1 = \emptyset$ and $S_2 = \{3\}$, i.e., agent one is *never* monitored and agent two is monitored in the low state. Pareto optimal contracts are given by $t_1 = -t_2 = (2+c, -2, 2+c, -2)$, since under this contract agent two is completely insured, i.e., consumption is state-independent (net of monitoring costs). However, agent one's net-transfer is not constant even though the agent is never monitored. Incentive compatibility for agent two is straightforward. Incentive compatibility for agent one is fulfilled in expected value: Assume that agent one gets the high realization. The expected

[29] To check this use the same argument as in step (ii) of the proof of Theorem 2, and take the expected value.

net transfer is $c/2$, the same expected net-transfer the agent would get in the low state. The argument goes through even if agent one is slightly risk averse because this arrangement economizes on monitoring costs: Choosing $S_1 = \{3\}$ increases monitoring costs by a discrete amount since monitoring is deterministic.

Some readers may be tempted to construe Example 1 as refuting the optimality of debt even under deterministic verification. We regard this interpretation as misguided. As Townsend (1987, p. 382) notes, the motivation for an analysis such as our Theorem 1 is to "begin with some striking arrangement [e.g, debt] in an actual economy and ask whether any theoretical environment might yield such an arrangement ... without making the [model] too complicated or implausible." We view the question—is a model with an expected value incentive constraint better than a model with a point-wise constraint?—to be methodologically equivalent to the question—is a model with stochastic monitoring better than a model with deterministic monitoring? In our opinion the answer is clearly no. Mathematical generality is not the desideratum per se, rather it is the consistency of the *structure and results* of alternative models with those observed in actual economic environments which determines which model is more appropriate for the problem at hand.

In fact, the appropriateness of the point-wise constraint appears to be directly linked to the appropriateness of deterministic monitoring. Both specifications seem to be consistent with key institutional features of US bankruptcy procedures.[30] When a firm petitions for bankruptcy protection under Chapter 7 of the US Bankruptcy Code, a trustee is appointed by the court. This trustee is bound by law to give a full account of the status of the claims owed by and to the insolvent firm, by every individual involved in transactions with it. Formally, this corresponds to the game with communication which leads to the (point-wise) incentive constraints IC1 and IC2. Thus, even though an insolvent firm's creditors are likely to have information about the firm's assets, the Bankruptcy Code prohibits them by law from attempting to secure direct payments from the firm or from those who owe payments to it. One interpretation of the co-existence of different institutions

[30]See White (1989) for a detailed discussion of the corporate bankruptcy decision in the US. Note that we take the procedures associated with Chapter 7 of the US Bankruptcy Code as given. An analysis of why these particular legal structures have emerged is beyond the scope of this paper.

is that there are different equity versus efficiency tradeoffs in bankruptcy and auditing problems. Perhaps society is willing to pay a higher price for fairness in bankruptcy settings because all agents are potentially subject to a random shock which could render them insolvent.

We conclude by discussing the Boyd and Prescott problem noted in the Introduction: the nature of optimal multilateral contracts in economies with additional information imperfections such as adverse selection and moral hazard. Boyd and Smith (1991) introduce adverse selection into the (deterministic) costly state verification model when contracts are restricted to be bilateral and agents are risk neutral. Agent heterogeneity is clearly essential for such problems, and our model permits agents to differ on several different dimensions (i.e., preferences, (endowment) distribution functions, transfer functions, and monitoring cost functions need not be identical). We believe that our multilateral results will be robust even when these additional imperfections are introduced, but this remains for future research.

6 Appendix

Proof of Remark 1. Let t be a simple function on Y_2, i.e., there exist $A_i \in \beta_2$, and $\lambda_i \in I\!R$ such that $t = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}$, where

$$\mathbf{1}_{A_i}(x) = \begin{cases} 1 & x \in A_i; \\ 0 & \text{otherwise.} \end{cases}$$

Then $t(g(x)) = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{g^{-1}A_i}(x)$. Hence,

$$\begin{aligned} \int_{Y_1} t(g(x))\, d\mu_1(x) &= \sum_{i=1}^{n} \lambda_i \int_{Y_1} \mathbf{1}_{g^{-1}A_i}(x)\, d\mu_1(x) = \sum_{i=1}^{n} \lambda_i \mu_1(g^{-1}A_i) \\ &= \sum_{i=1}^{n} \lambda_i \mu_2(A) = \sum_{i=1}^{n} \lambda_i \int_{Y_2} \mathbf{1}_{A_i}(x)\, d\mu_2(x) = \int_{Y_2} t(x)\, d\mu_2(x). \end{aligned}$$

The third inequality follows because g is measure preserving. Since the Remark holds for all simple functions, it also holds for all integrable functions.[31]

Lemma 1. *Let μ be a measure on $I\!R$ and let A, B be two subsets of $I\!R$ with the same measure. Let f be an integrable function on $I\!R$. Assume that*

[31]This is a standard approximation argument in measure theory: All integrable functions can be approximated by simple functions.

$a < b$ for every $a \in A$ and for every $b \in B$. Assume that f is bounded on $A \cup B$. Let g be a measure preserving isomorphism on $I\!R$ such that $g(x) = x$ for every $x \in I\!R \setminus A \cup B$. Assume that there exist γ_A, $\gamma_B \in I\!R$ such that $x + f(x) \leq \gamma_A \leq x + f(g(x))$, for every $x \in A$; and $x + f(g(x)) \leq \gamma_B \leq x + f(x)$, for every $x \in B$. Then $x + f(g(x))$ is less risky than $x + f(x)$ in the Rothschild and Stiglitz sense (i.e., every risk averse agent prefers $x + f(g(x))$ over $x + f(x)$).

Proof. Here we need only check that the integral condition of Rothschild and Stiglitz (1970) holds. Let F be the distribution of $x + f(g(x))$ and let G be the distribution of $x + f(x)$. Then
(i) $G(t) - F(t) \geq 0$ for every $t < \gamma_A$.
(ii) $G(t) - F(t)$ is monotonically decreasing for $\gamma_A \leq t < \gamma_B$; and
(iii) $G(t) - F(t) \leq 0$ for every $t \geq \gamma_B$.
(i) follows from the fact that g is measure preserving. In particular:

$$\mu\left(\{x \in A: x + f(x) \leq t\}\right) = \mu\left(\{x \in B: g(x) + f(g(x)) \leq t\}\right).$$

Note that $\mu\left(\{x \in B: x + f(x) \leq t\}\right) = 0$, and $\mu\left(\{x \in A: x + f(g(x)) \leq t\}\right) = 0$ for every $t < \gamma_A$. Since $g(x) \leq x$ for every $x \in B$ this proves (i). (ii) follows immediately from the definition of γ_A and γ_B. Finally, the argument for (iii) is similar to the argument for (i).

Since f is bounded and since A and B are bounded there exists an $M > 0$ such that $G(t) - F(t) = 0$ for every $t \notin [-M, M]$.[32] Let $T(y) = \int_{-M}^{y} G(t) - F(t)\, dt$. By Rothschild and Stiglitz (1970, Theorem 2) it is sufficient to prove that the following two conditions are satisfied.
(a) $T(M) = \int_{-M}^{M} [G_i(x) - F_i(x)]\, dx = 0$;
(b) $T(y) \geq 0$ for $-M \leq y \leq M$.
(a) follows immediately from integration by parts and from the fact that g is measure preserving.[33] (b) follows immediately from (a) and from conditions

[32]In order to be able to apply the Theorem we need that the points of increase of the distribution functions lie in a compact interval.

[33]Measure preservingness implies $\int_{-M}^{M} t\, dF(t) = \int_{-M}^{M} t\, dG(t)$. Partial integration therefore yields

$$\int_{-M}^{M} G(t) - F(t)\, dt = t(G(t) - F(t))|_{-M}^{M} - \left(\int_{-M}^{M} t\, dG(t) - \int_{-M}^{M} t\, dF(t)\right) = 0.$$

(i), (ii) and (iii). This concludes the proof.

Lemma 2. *Let u be a utility function which is twice continuously differentiable. Assume that $u''(x) < 0$ for every x. Let A, B be two subsets of $\mathbb{R}$ with the same measure. Let f be integrable and let g be a measure preserving transformation such that $g(g(x)) = x$ for every x, such that $g(A) = B$, and $f(x) = x$ for every $x \notin A \cup B$. Assume that $f(a) < f(b)$ for every $a \in A$ and for every $b \in B$. Then Lemma 1 holds with a strict inequality, i.e., the agent strictly prefers the contract $x + f(g(x))$ to $x + f(x)$.*

Proof. Here we need only check that the integral condition of Rothschild and Stiglitz holds with a strict inequality, and then use partial integration to show that the agent strictly prefers $x+f(g(x))$ (cf., Rothschild and Stiglitz (1970), footnote 10).

Let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that $f(b) - f(a) > \delta$ except on sets $S_A \subset A$ and $S_B \subset B$ with $\mu(S_A) = \mu(S_B) < \varepsilon$. Since $g(g(x)) = x$ for every $x \in \mathbb{R}$, we can construct a finite partition A_i, $i = 1, \ldots, n$ of $A \setminus S_A$ and a finite partition B_i, $i = 1, \ldots, n$ of $B \setminus S_B$ such that $g(A_i) = B_i$ and such that the condition of Lemma 1 is fulfilled for each A_i and B_i.[34] Thus, since we can subsequently exchange the transfers of A_i with the transfers of B_i and since ε was chosen arbitrarily, Lemma 1 implies that $x + f(g(x))$ is less risky than $x + f(x)$. Thus, the integral conditions of Rothschild and Stiglitz (1970) hold. Note that $T(y) > 0$ on a set of positive measure.

Integration by parts yields

$$\begin{aligned} \int_{-M}^{M} u(x)\,dS(x) &= u(x)S(x)|_{-M}^{M} - \int_{-M}^{M} u'(x)S(x)\,dx \\ &= -\,u'(x)T(x)|_{-M}^{M} + \int_{-M}^{M} u''(x)T(x)\,dx, \end{aligned} \tag{A.1}$$

since $u(x)S(x)|_{-M}^{M} = 0$ by (a). Further, since T is strictly positive on a set of positive measure, and since $u'' < 0$ it follows that

$$\int_{-M}^{M} u''(x)T(x)\,dx < 0. \tag{A.2}$$

[34]That means that there exist γ_{A_i}, γ_{B_i} for $i = 1, \ldots, n$ such that $x + f(x) \le \gamma_{A_i} \le x + f(g(x))$, for every $x \in A_i$; and $x + f(g(x)) \le \gamma_{B_i} \le x + f(x)$, for every $x \in B_i$.

(A.1), (A.2) and $u'(x)T(x)|_{-M}^{M} = 0$ immediately imply that the agent's utility is strictly greater under contract $x + f(g(x))$. This proves the Lemma.

Next we state a "generalized" version of the Isomorphism Theorem. The problem we face in the proof of Theorem 2 is that the sets $\mathcal{U}$ and $\mathcal{V}$ are not necessarily representable as the product of lower dimensional spaces. However, for the proof of the Theorem we need an isomorphism between subsets $A \subset \mathcal{U}$ and $B \subset \mathcal{V}$ which is also an isomorphism between the corresponding "slices" of A and B. The existence of such an isomorphism and of the subsets is provided by the following Lemma. The central step of the argument is the use of a theorem on measurable selections.

Lemma 3. *Let K^i, $i = 1, 2$ be two compact subsets of $\mathbb{R} \times \mathbb{R}^n$. Let μ_1 and μ_n be probability measures on $\mathbb{R}$ and $\mathbb{R}^n$, respectively, and let μ denote the product measure. Then there exist measurable subsets $A^i \subset K^i$ and measure preserving mappings $h_1: A^1 \to A^2$ and $h_2: A^2 \to A^1$ such that $x \mapsto h_i(x, y)$, $i = 1, 2$ are measure preserving mappings from A^1_y to A^2_y and from A^2_y to A^1_y, respectively, for every $y \in \mathbb{R}^n$ where $A^i_y = \{(x, y) : (x, y) \in K^i\}$.*[35]

Proof. Define a function $f: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}_+$ by

$$f(x, y, t) = \left|\mu_1\left((-\infty, x) \times \{y\} \cap K^1_y\right) - \mu_1\left((-\infty, x + t) \times \{y\} \cap K^2_y\right)\right|.$$

Note that f is jointly measurable in x and y.[36] Furthermore, for fixed x and y, the function $t \mapsto f(x, y, t)$ is continuous on K_1. By compactness of K_1 and K_2 there always exists a $\bar{t}$ such that $f(x, y, \bar{t}) = \inf_t f(x, y, t)$. Thus, Landers' Theorem [cf., 6.10 of Strasser (1985)][37] implies that there exists a Borel measurable function $\phi: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x, y, \phi(x, y)) = \inf_t f(x, y, t).$$

[35]Let B be a subset of A^i_y. Then $B = B' \times \{y\}$. By slight abuse of notation, we define $\mu_1(B)$ to be $\mu_1(B')$.

[36]This can be established as follows: Let $g: K_1 \to \mathbb{R}$ be defined by $(x, y) \mapsto \mu_1\left((-\infty, x) \times \{y\} \cap K^1_y\right)$. For fixed x, the mapping $y \mapsto g(x, y)$ is measurable by Fubini's Theorem. Furthermore, note that for fixed y the mapping $x \mapsto g(x, y)$ is continuous. Thus, g is jointly measurable [see Castaing and Valadier (1977)]. A similar argument shows that $(x, y, t) \mapsto \mu_1\left((-\infty, x + t) \times \{y\} \cap K^2_y\right)$ is also jointly measurable in x and y. This proves the measurability of f in x and y.

[37]Landers' Theorem: Let $(\Omega, \mathcal{A}, \mu)$ be a measurable space and (X, d) a σ-compact metric space. Let $h: \Omega \times X \to \mathbb{R}$ be a function such that $\omega \mapsto h(\omega, x)$ is measurable for every

Note that ϕ is part of the measure preserving transformation defined below, however, we still must construct the sets A^i, i.e., the sets where the measure of the slices coincide. We do this as follows. Let

$$A^1 = \{(x,y)\colon f(x,y,\phi(x,y)) = 0\}.$$

Then A^1 is measurable. Define h_1 on A^1 by $(x,y) \mapsto (x+\phi(x,y),y)$. In a similar way we can construct a measurable subset A^2 of K^2 and a mapping h_2 on A_2. It immediately follows from the above construction that $h_1(A^1_y) \subset A^2_y$ and $h_2(A^2_y) \subset A^1_y$. Thus, $h_1(A^1) \subset A^2$ and $h_2(A^2) \subset A^1$. It now remains to show that the h_i are measure preserving. This can be established as follows:

Let $y \in I\!R^n$. By construction, h_i preserves the μ_1-measure of all sets of the form $(-\infty,a) \times \{y\} \cap A^i_y$ for $i = 1,2$. These sets generate the σ-algebra of all measurable sets. Thus h_i is a measure preserving transformation on K^i_y for $i = 1,2$. Finally, note that Fubini's Theorem proves that the mappings h_i are measure preserving on K^i. This concludes the proof of the Lemma.

$x \in S$ and $t \mapsto h(\omega,t)$ is continuous for every $\omega \in \Omega$. If

$$B(\omega) = \left\{x \in X\colon h(\omega,x) = \inf_{y\in X} h(\omega,y)\right\} \neq \emptyset$$

for a.e. ω then there is a measurable function $\phi\colon \Omega \to X$ such that

$$h(\omega,\phi(\omega)) = \inf\{h(\omega,y)\colon y \in X\}\,,\ \mu\text{-a.e.}$$

References

R. Ash (1972), "Real Analysis and Probability," Academic Press: New York.

K. Border and J. Sobel (1987), "Samurai Accountant: A Theory of Auditing and Plunder," *Review of Economic Studies* **54**, 525–540.

J. Boyd and E.C. Prescott (1986), "Financial Intermediary Coalitions," *Journal of Economic Theory* **38**, 211–232.

J. Boyd and E.C. Prescott (1987), "Dynamic Coalitions, Growth, and the Firm," E.C. Prescott and N. Wallace (eds.), Contractual Arrangements for Intertemporal Trade, University of Minnesota Press.

J. Boyd and B. Smith (1991), "Intermediation and the Efficient Allocation of Investment Capital in the Presence of Adverse Selection and Costly State Verification," mimeo, Federal Reserve Bank of Minneapolis.

Y.M. Caspi (1978), "A Limit Theorem on the Core of an Economy with Individual Risks," *Review of Economic Studies* **45**, 267–271.

C. Castaing and M. Valadier (1977), "Convex Analysis and Measurable Multifunctions," Lecture Notes in Mathematics, Springer Verlag: New York.

D. Gale and M. Hellwig (1985), "Incentive–Compatible Debt Contracts: The One Period Problem," *Review of Economic Studies* **52**, 647–663.

J. Harsanyi (1967), "Games with Incomplete Information Played by Bayesian Players," *Management Science* **14**, 159–182.

S. Krasa and A. Villamil (1991a), "Monitoring the Monitor: An Incentive Structure for a Financial Intermediary," forthcoming *Journal of Economic Theory*.

S. Krasa and A. Villamil (1991b), "Existence of Optimal Contracts in the Costly State Verification Model," mimeo, University of Illinois.

D. Mookherjee and I. Png (1989), "Optimal Auditing, Insurance and Redistribution," *Quarterly Journal of Economics* **104**, 399-415.

K. Parthasarathy (1977), "Introduction to Probability and Measure," McMillan Press: London.

E.C. Prescott and R.M. Townsend (1984), "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica* **52**, 21–45.

M. Rothschild and J. Stiglitz (1970), "Increasing Risk: I. A Definition," *Journal of Economic Theory* **2**, 225–243.

H. Strasser (1985), "Mathematical Theory of Statistics," de Gruyter: Berlin, New York.

R.M. Townsend (1979), "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory* **21**, 1–29.

R.M. TOWNSEND (1987), "Arrow-Debreu Programs as Microfoundations of Macroeconomics," T. Bewley (ed.), *Advances in Economic Theory*, Econometric Society Monograph, Cambridge University Press.

R.M. TOWNSEND (1988), "Information Constrained Insurance: The Revelation Principle Extended," *Journal of Monetary Economics* **21**, 411-450.

M. WHITE (1989), "The Corporate Bankruptcy Decision," *Journal of Economic Perspectives* **3**, 129–151.

S.D. WILLIAMSON (1986), "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing," *Journal of Monetary Economics* **18**, 159–179.

R. WILSON (1968), "On the Theory of Syndicates," *Econometrica* **36**, 119–132.

A. WINTON (1991), "Costly State Verification and Multiple Investors: The Role of Seniority," mimeo, Northwestern University.