

OPTIMAL AUDITING, INSURANCE, AND REDISTRIBUTION*

DILIP MOOKHERJEE AND IVAN PNG

I. INTRODUCTION

In both market and planned economies, organizations rely on communication of information about realized levels of performance or states of nature. For instance, lenders and insurers rely on clients' reports of income; central planning bureaus rely on performance reports from productive enterprises under their direction; and revenue collection agencies rely on income reports from taxpayers.

In all these examples, the reporting agents have an incentive to misrepresent because transfers between the parties depend on their reports. In practice, an important tool of the recipient of reports (the principal) is the right to audit the information communicated by the agent. Such audits usually entail significant costs. Consequently, the principal would generally prefer to audit the agents' reports selectively. This gives rise to the question: what is the optimal nature of audit strategies, and how are these related to the structure of optimal transfers?

In an important paper Townsend [1979] analyzed the optimal insurance contract between a risk-averse agent with private information about the realization of his income, and an uninformed risk-neutral principal who had the right to audit the agent's report of (realized) income at a fixed cost. Townsend restricted attention to the class of *deterministic* audit strategies, where the principal can audit reports only with probability either one or zero. Within this class he showed that the optimal strategy involved auditing all income reports below a certain cutoff level with probability one, and not auditing any higher income reports. Consequently, under the optimal scheme, the agent must pay a constant amount to the principal in all states above the cutoff, and receive a constant consumption level in all states below the cutoff. These contracts

*We thank David Baron, David Besanko, Joel Demski, Jeffrey Lacker, Nahum Melumad, Joel Sobel, Mark Wolfson, and the reviewers for helpful comments; and Glenn Graves, Dan Asquith, and Octavio Colmenares for assistance with numerical solutions. The second author gratefully acknowledges support of a grant from the Academic Senate of the University of California.

thus resemble debt contracts, provided that the act of auditing is identified with bankruptcy.¹

Townsend noted, however, that the optimality of such contracts might depend crucially on the assumption of deterministic auditing. By way of a simple example with two possible income levels, he showed that the optimal contract with deterministic auditing may be dominated by one with random audits. But he left open the question of the general character of optimal contracts when random audits are permitted.

The objective of this paper is to analyze this question. Our results should be relevant to the design of optimal verification strategies in insurance or tax contexts. We are interested particularly in finding general conditions under which random audits are optimal. We also aim to analyze the optimality of debt contracts, thus permitting us to examine the robustness of Townsend's results in the presence of random audits.

In addition to allowing random auditing, our model generalizes Townsend's by allowing the likelihood of alternative income states to depend on actions chosen by the agent that cannot be observed by the principal at any cost. Our formulation of moral hazard extends the model of Grossman and Hart [1983]. Thus, we assume that the agent first chooses an unobservable action, following which a certain income level is realized. The agent privately observes his realized income and reports it to the principal, who may subsequently audit the report at a cost. The ability of the principal to offer insurance to the agent is constrained both by unobservability of the agent's action, as well as imperfect observability of the realization of income.

In addition to considering the issue of moral hazard, our analysis differs from Reinganum and Wilde [1985] and Border and Sobel [1987] in two essential respects. In both models the principal's objective is to maximize net revenues collected, without any participation constraint on the expected utility of the agent. Second, both assume that the agent is risk-neutral. As Border and Sobel demonstrate, an optimal policy then typically will not exist as the agent may be induced to report truthfully with an arbitrarily

1. A debt contract requires the agent to pay a constant amount to the principal. If realized income falls short of the required payment, the agent declares bankruptcy; in such a case all assets are transferred to the principal; and the agent's consumption is fixed by limits on his liability. See the concluding section for further elaboration.

Gale and Hellwig [1985] study an extension of this model, where principal and agent additionally choose a level of investment, which is mutually observed.

small deadweight loss (by a scheme that audits the agent's report with very low probability and provides correspondingly large rewards for reports that are verified to be truthful). Consequently, Border and Sobel impose an arbitrary upper bound on the size of rewards that may be paid for reports that are verified to be truthful.

We show that provided the agent has a minimal degree of risk aversion, specifically that he is not subject to the classical St. Petersburg paradox, optimal contracts exist in the absence of any upper bound on consumption or rewards. Our main result is that *all* audits must be random in any optimal contract satisfying the weak condition that the agent obtains positive consumption in every realization of income. The intuition for this is that as long as there are nontrivial penalties for those discovered to have reported falsely, an audit probability equal to one ensures that the agent will strictly prefer not to lie. Consequently, the probability of audit can be reduced somewhat without distorting the incentives of the agent for choosing appropriate actions or to report truthfully. This achieves a reduction in expected audit costs, which in turn permits an increase in the expected utility of the agent.

Given that contracts are random, the question then arises: should those audited and verified to have reported truthfully be given the same level of consumption as those with the same realized income but not audited? We show that it is always optimal to reward agents who are verified to have reported truthfully. Thus, if the agent's realized income is one for which auditing is prescribed, and if the agent reports truthfully, he will bear risk associated with the decision of the principal to audit his report. Specifically, the agent strictly prefers to be audited.

Finally, we investigate monotonicity properties of optimal contracts. If there are only two possible income levels, the optimal contract will require that the high income report not be audited, and also be associated with a higher transfer, and higher consumption and utility, while the low income report must be audited with some probability. But if there are three possible income levels, consumption and probability of audit need not be monotone. Even in the absence of moral hazard, it is possible for the agent's consumption, conditional on not being audited, to be nonmonotone in his realized income. In contrast, Border and Sobel [1987] show that the optimal policy in the model with a risk-neutral agent and no moral hazard provides consumption and utility that is monotone in income.

The subject of this paper is related also to the work of Baiman and Demski [1980] and Dye [1986] who analyzed optimal contracts between a risk-cum-effort-averse agent, and a risk-neutral principal. In their setting, both agent and principal freely observe realized income, and on the basis of observed income, the principal could go back to investigate the agent's effort at a cost. These models involve investigations of effort levels, in contrast to our model which involves verification of agents' income reports. Thus, in their model, investigation strategies are designed to improve the agent's effort incentives, whereas in our model, verification affects both the incentive for effort incentives and the incentives to report truthfully. The results of the two models are also quite different: Dye presents a variety of sufficient cases where optimal investigation probabilities are zero or one.

The next section of the paper presents our model and results. The concluding section summarizes implications and describes extensions of our analysis.

II. MODEL AND RESULTS

Following Grossman and Hart [1983], we assume that there are n possible levels of realized income: Y_1, Y_2, \dots, Y_n , where $Y_{i-1} < Y_i$ for all i . The agent must choose an action from some set A of possible actions. The action chosen determines the probability distribution over levels of realized income: $\lambda_i(a)$ denotes the probability of income Y_i resulting from action a , where $\lambda_i(a) \geq 0$ and $\sum_{i=1}^n \lambda_i(a) = 1$.

The principal is risk-neutral. The agent is risk-averse, and has a von Neumann-Morgenstern utility function W defined over his consumption and action. We assume

- A1. W is additively separable in consumption and action, $W = U(C) - G(a)$.² U is strictly increasing, differentiable, and concave. Further, the agent's consumption must be nonnegative, and $U(0) = 0$.

This assumption bounds the extent to which the agent can be punished, and thereby rules out schemes that may attain approximately first-best outcomes by imposing arbitrarily large penalties for false reporting or for undesirable realizations of income. There

2. Our results also extend to the context of multiplicative separability of the agent's utility, using very similar arguments.

remains, however, the possibility that the principal may encourage truthful reporting at minimal cost by auditing with extremely low probability, and simultaneously offering unboundedly large rewards for verification of truthful reporting.³ To rule this out, we assume a minimal degree of risk aversion, specifically, that he is not subject to the classical St. Petersburg Paradox:

$$A2. \lim_{c \rightarrow \infty} (U(C))/C = 0.$$

Utility functions in the constant absolute risk aversion and constant relative risk aversion classes satisfy this condition. We show below that this assumption ensures that optimal schemes exist.

The agent privately observes his realized income and must report his income to the principal. The principal may discover the agent's true realized income Y_i , but the process of audit involves a positive cost E_i (which may depend on the true income Y_i).⁴ By extending the argument in Melumad and Mookherjee [1989] to the case of moral hazard, it may be verified that the Revelation Principle applies to this setting. Therefore, to solve for an optimal scheme, the principal can restrict attention to schemes where the agent is given incentives to report his realized income truthfully, as well as choose appropriate actions.

Consider an agent with realized income Y_i . If the agent reports Y_l , he must pay a transfer T_l to the principal. The principal audits the report with probability p_l . If there is an audit, the agent's true income Y_i will be discovered, and the agent must pay a penalty F_{li} . The objective of the principal is to find the scheme that maximizes the expected utility of the agent, subject to a minimum expected profit constraint, and subject to the appropriate incentive constraints.⁵

As a preliminary, notice that in equilibrium, the agent will report all realizations of income truthfully. Thus, to provide maximum incentive to report truthfully without affecting the agent's expected utility, the penalties F_{li} levied when a report is found to be false, $l \neq i$, can be made as large as possible. So if the agent is detected to have lied, his consumption can be set equal to zero: $Y_i - T_l - F_{li} = 0$ if $l \neq i$. To simplify notation, let F_{li} be denoted by F_i .

3. If the agent is risk-neutral, an optimal scheme does not exist (see Border and Sobel [1987]).

4. The agent may also incur costs in the process of being audited. However, transfers to the principal can be interpreted to include these costs, so we can without loss of generality assume that all audit costs are borne by the principal.

5. We assume that transfers are deterministic. In the discussion following Proposition 1 we indicate how the results of the proposition extend to the case where randomized transfers are permissible.

The principal's problem is then to find a scheme $(\{p_i\}, \{T_i\}, \{F_i\})$ together with an appropriate action choice a for the agent to maximize the expected utility of the agent

$$(1) \quad \sum_{i=1}^n \lambda_i(a) [p_i U(Y_i - T_i - F_i) + (1 - p_i) U(Y_i - T_i)] - G(a),$$

subject to the following constraints:

$$(2) \quad V_i \equiv p_i U(Y_i - T_i - F_i) + (1 - p_i) U(Y_i - T_i) \\ \geq (1 - p_h) U(Y_i - T_h) \quad \text{for all } i, h,$$

$$(3) \quad \sum_{i=1}^n \lambda_i(a) V_i - G(a) \geq \sum_{i=1}^n \lambda_i(a') V_i - G(a') \quad \text{for all } a' \in A,$$

$$(4) \quad \sum_{i=1}^n \lambda_i(a) [T_i - p_i(E_i - F_i)] \geq R,$$

$$(5) \quad Y_i - T_i - F_i \geq 0, \quad Y_i - T_i \geq 0, \quad 1 \geq p_i \geq 0, \quad \text{for all } i.$$

Inequalities (2) represent the *reporting constraints* that require the agent to always prefer to report his realized income truthfully. Inequalities (3) represent the *effort constraints* that require the agent to prefer the choice of action recommended by the principal to any other. Inequality (4) requires the expected profit of the principal to meet some minimum preassigned level R . Any scheme that satisfies (2)–(5) will be called *feasible*. Henceforth, we shall assume that the parameters of the problem are such that there exists at least one feasible solution.

To interpret the problem in the context of taxation, suppose that the economy consists of a large population of ex ante identical taxpayers. The ex post income of each taxpayer is independently drawn from a distribution, represented by $\lambda_i(a)$, that depends on individual efforts a . Let the population be large enough to permit application of the law of large numbers; then given a common level of effort, a , the ex post distribution of income will be given by $\lambda_i(a)$. In this interpretation (1) represents a utilitarian welfare function for a social planner, which is to be maximized by choice of taxes, audit probabilities, and penalties, subject to the incentive constraints (2) and (3) for taxpayers, a minimum revenue constraint (4) for the planner, and constraints that consumption be nonnegative and that the p_i be probabilities (5).

We have formulated the problem in a way traditionally employed to analyze competitive insurance markets, i.e., to maximize the expected utility of the agent subject to a constraint of

minimum expected profit for the principal. The following lemma implies that this formulation is equivalent to the dual problem of maximizing the principal's expected profit subject to an ex ante participation constraint for the agent.

LEMMA 1. Any scheme in which the expected profit constraint (4) does not bind is not optimal.

Proof of Lemma 1. Suppose otherwise: that is, there is an optimal scheme $(\{p_i\}, \{T_i\}, \{F_i\})$ together with an appropriate action choice a , such that the expected profit constraint is not binding. Let I denote the set of incomes Y_i for which $p_i > 0$, and J the set of remaining incomes. Consider a variation that differs from the original scheme only in the following regard. Given $\delta > 0$, for each Y_i in I , replace F_i by $F_i(\delta)$, which is defined by $p_i U(Y_i - T_i - F_i(\delta)) + (1 - p_i)U(Y_i - T_i) = V_i + \delta$. Clearly, $F_i(\delta)$ is continuous in δ and satisfies $F_i(0) = F_i$. For each Y_j in J let there be a reward of an arbitrary amount $-\hat{F}_j > 0$ for verification of a truthful report, and let the probability of audit be $p_j(\delta)$, where $p_j(\delta)$ is defined by $p_j(\delta) U(Y_j - T_j - \hat{F}_j) + [1 - p_j(\delta)] U(Y_j - T_j) = V_j + \delta$. Then $p_j(\delta)$ is continuous in δ , with $p_j(0) = 0$.

In this variation the reporting constraints continue to be satisfied, and since expected utility in every realization of income is increased by an identical amount δ , the effort constraints are not affected. The loss in expected profits associated with this variation is

$$L(\delta) = \sum_{i \in I} \lambda_i(a) p_i [F_i - F_i(\delta)] + \sum_{j \in J} \lambda_j(a) p_j(\delta) [E_j - \hat{F}_j],$$

which is continuous and satisfies $L(0) = 0$. Hence, for any given slack D in the budget constraint (4), there exists an amount δ by which expected utility can be increased in every realization of income, at a reduction of profit less than D . Thus, the original scheme is not optimal.

Q.E.D.

The next Lemma shows that the principal should not penalize the agent if the agent's report is audited and found to be truthful.

LEMMA 2. Without loss of generality, the principal should not penalize the agent if the agent's report is verified to be truthful; that is, $F_i \leq 0$ for all i .

Proof of Lemma 2. Suppose that there exists a feasible scheme with $F_i > 0$. Then the principal can lower F_i to 0 and simultaneously

increase T_i , in such a way as not to affect the agent's expected utility V_i for realized income Y_i . By construction, then, the reporting and effort constraints are preserved. But the agent no longer bears risk with income Y_i , so the principal's expected profits cannot decrease, and, consequently, the new scheme is feasible.

Q.E.D.

For those realized incomes where $p_i = 0$ or 1, the value of F_i is arbitrary, and hence may be set equal to zero. We are now in a position to state and prove the main result. For a large class of utility functions—all those with the property that marginal utility at zero consumption is infinitely large, the optimal scheme will provide positive consumption in every state. This class includes utility functions with constant relative risk aversion. Proposition 1 shows that, for utility functions in this class, any income report that is audited must be audited randomly. More generally, any particular income report must be audited randomly, as long as all other income levels associated with a higher transfer allow the agent a positive level of consumption.

PROPOSITION 1.

- a. Optimal schemes exist.
- b. All income reports that are audited must be audited randomly—that is, $p_i < 1$ for all i —in any optimal scheme that provides the agent positive consumption in every income realization.
- c. Every optimal scheme has the property that if the agent's report is audited and verified to be truthful, the agent must be rewarded; that is, $F_i < 0$ whenever $1 > p_i > 0$.

Proof of Proposition 1. a. Given any feasible scheme s , let the corresponding level of expected utility be $\phi(s)$. It suffices to show that for any sequence $\{s^v\}$ of feasible schemes with the property that $\phi(s^v) \rightarrow \phi^*$, there exists a feasible scheme s^* such that $\phi(s^*) = \phi^*$. Let \hat{s} be a limit point of $\{s^v\}$; then by continuity of $U, \phi(\hat{s}) = \phi^*$. If $\hat{F}_i > -\infty$ for all i , then \hat{s} is clearly feasible because taxes and audit probabilities are uniformly bounded. So suppose instead that $\hat{F}_i = -\infty$ for some i ; that is, \hat{s} is infeasible.

Let λ_i^v denote $\lambda_i(a^v)$. Using (4), s^v is feasible implies that $\sum_i \lambda_i^v p_i^v F_i^v \geq R - \sum_i \lambda_i^v T_i^v$. Since $T_i^v \leq Y_i \leq Y_n$, we have upon using Lemma 2 $\lambda_i^v p_i^v F_i^v \geq R - Y_n$ for all i , implying that

$$0 \leq \lambda_i^v p_i^v U(Y_i - T_i^v - F_i^v) \leq ((Y_n - R)/-F_i^v) U(Y_i - T_i^v - F_i^v).$$

By (A2), $\hat{F}_i = -\infty$ then implies that $\lambda_i^v p_i^v U(Y_i - T_i^v - F_i^v) \rightarrow 0$.

Consider the scheme s^* defined by $p_i^* = \hat{p}_i$, $T_i^* = \hat{T}_i$, $a^* = \hat{a}$ and

$$F_i^* = \begin{cases} 0 & \text{if } \hat{F}_i = -\infty \\ \hat{F}_i & \text{if } \hat{F}_i > -\infty. \end{cases}$$

It may be checked that s^* is feasible and yields $\phi(s^*) = \phi^*$.

b. Consider any scheme where some income report Y_i is always audited; that is, $p_i = 1$. Then if the agent's true income is Y_j , $j \neq i$, a report of Y_i will lead to zero utility. If the scheme offers positive utility for a truthful report of income Y_i , the agent will then strictly prefer not to report Y_i . Now $p_i = 1$ implies that $F_i = 0$; hence a reduction in p_i does not affect V_i . Thus, p_i can be lowered slightly, without disturbing any reporting constraints or effort constraints. This lowers the expected cost of audit and thus generates additional expected profits. By Lemma 1, the original scheme could not be optimal.⁶

c. Suppose that there is an optimal scheme for which there exists i such that $1 > p_i > 0$ and $F_i = 0$. Consider the following variation on this scheme. Let T_i be increased slightly, $dT_i > 0$, and the reward for truthful reports of income Y_i also be increased slightly, $dF_i < 0$, so that the expected income of the agent in state i is unaffected: $dT_i + p_i dF_i = 0$. Since initially, $F_i = 0$ —that is, the agent with income Y_i was perfectly insured under the original scheme—the variation imposes no first-order effect on expected utility: $dV_i = 0$. But now that T_i has been increased, we can reduce p_i slightly without disturbing any reporting constraints or effort constraints. This enables a first-order reduction in the expected cost of audit; so by Lemma 1, the original scheme could not have been optimal.

Q.E.D.

In Proposition 1 we confined attention to schemes with non-randomized transfers. In the income tax context this may be justified by a requirement of horizontal equity. Proposition 1 generalizes to the case of randomized transfers as well. The existence of an optimal scheme is straightforward if we restrict the consumption of the agent to a finite number of alternative values, as in Prescott and Townsend [1984] and Townsend [1988]. Proposition 1b generalizes to the statement that all audits must be random, provided that the optimal scheme does not provide the agent with

6. This argument is essentially similar to that given by Townsend [1979] in his example in which the optimal scheme with deterministic verification was dominated by one with stochastic verification.

zero consumption with probability one at any income. The generalization of Proposition 1c is that every optimal scheme has the property that following a truthful report, the expected utility of the agent subsequent to an audit but prior to the randomization of the associated transfer is higher than the corresponding expected utility if the agent were not audited. These results also extend to the case where effort may be randomized as well. All these extensions follow from straightforward adaptation of the arguments for the deterministic case.

We now turn to analyze the relation between the optimal audit probabilities and optimal transfers. We first establish a lemma that expresses the notion that the purpose of auditing is to deter strategic misreporting by the agent.

LEMMA 3. In any optimal scheme if an income report Y_i is audited with positive probability, there must exist an income Y_j such that the agent with true income Y_j is indifferent between reporting truthfully and reporting Y_i .

Proof of Lemma 3. Suppose otherwise, that $p_i > 0$, while every other type strictly prefers to report truthfully rather than Y_i . If $p_i = 1$ and thus $F_i = 0$, we can reduce p_i slightly without affecting V_i , and without disturbing any reporting or effort constraints. The resulting reduction in audit costs implies via Lemma 1 that the scheme cannot be optimal.

If $p_i < 1$ and thus by Proposition 1c, $F_i < 0$, we can reduce consumption when audited, $dF_i > 0$, and increase consumption when not audited, $dT_i < 0$, in income state Y_i in such a way as to keep V_i as well as all reporting constraints unchanged. Since the agent bears less risk in state Y_i , this enables an increase in expected profits, and therefore, by Lemma 1, the scheme is not optimal.

Q.E.D.

PROPOSITION 2. In any optimal scheme,

- a. reports corresponding to the highest transfer will not be audited; all other reports must be audited with positive probability;
- b. reports corresponding to higher transfers will be audited with equal or lower probability; further, if $T_k > T_l$ and $1 > p_l$, then $p_k < p_l$.

Proof of Proposition 2. a. Suppose that $T_i \geq T_j$ for all j , and yet $p_i > 0$. Consider any other possible income Y_k . Since $F_k \leq 0$ and $T_k \leq T_i$, the expected utility of an agent with income Y_k who reports

truthfully, $V_k \geq U(Y_k - T_i) > (1 - p_i)U(Y_k - T_i)$. Hence, in every other income state Y_k , the agent strictly prefers not to report Y_i , which contradicts Lemma 3.

Consider some income Y_j with $T_j < T_i$. Then $U(Y_i - T_j) > U(Y_i - T_i) = V_i$, since $p_i = 0$. Hence, if $p_j = 0$, agents with income Y_i corresponding to the highest transfer will prefer to report Y_j ; that is, the reporting constraints are violated.

b. Suppose that there are incomes Y_k, Y_l with $T_k > T_l$, and either (i) $p_k > p_l$, or (ii) $p_k \geq p_l$ with $p_l < 1$. Then in any income Y_i distinct from Y_k , the agent strictly prefers to report Y_l rather than Y_k . Now from (2a), $p_l > 0$. Hence $p_k > 0$, and a slight reduction in p_k is possible, which is a contradiction.

Q.E.D.

Proposition 2 implies that any nontrivial scheme—that is, not involving $T_i = R, p_i = F_i = 0$ for all i —will have the property that conditional on not being audited, the transfer from the agent to the principal *will* depend on the agent's realized income. The reason is that in any nontrivial scheme, there must exist incomes Y_i and Y_j with $T_i < T_j$, and Propositions 1 and 2 then imply that $1 > p_i > 0$ and $p_j < p_i$. So in both realizations Y_i and Y_j , the probability that the agent will not be audited is positive. But in the event of no audit, the transfer to the principal will be T_i for the agent with income Y_i and T_j for the agent with income Y_j . Hence, if we identify the act of auditing with forcing the agent into bankruptcy, the agent does not pay a constant transfer across all states of solvency.

Proposition 2 also implies that audit probabilities are decreasing in reported income if and only if the optimal transfers are increasing. We now investigate conditions under which the optimal contract requires transfers and consumption to be increasing in income.

PROPOSITION 3. Suppose that there are only two possible realizations of income: Y_1 and Y_2 , and $Y_1 < Y_2$. Then in any optimal contract the high-income report will not be audited, and will be associated with a larger transfer $T_2 \geq T_1$, as well as higher expected utility and consumption in the event of no audit, $V_2 \geq V_1$ and $Y_2 - T_2 \geq Y_1 - T_1$.

Proof of Proposition 3. We first prove that $T_2 \geq T_1$. Suppose otherwise—that $T_2 < T_1$. Then by Proposition 2, $p_1 = 0$, while $p_2 > 0$. We claim that such a scheme is dominated by the trivial scheme where $T_1 = T_2 = R, p_1 = p_2 = 0$. Let \bar{a} be the action chosen by the agent in the optimal scheme. Then the expected utility of the agent

under the optimal scheme is

$$\lambda_1(\bar{a})U(Y_1 - T_1) + \lambda_2(\bar{a}) [p_2U(Y_2 - T_2 - F_2) + (1 - p_2)U(Y_2 - T_2)] - G(\bar{a}),$$

where $\lambda_1(\bar{a})T_1 + \lambda_2(\bar{a})[T_2 - p_2(E_2 - F_2)] \geq R$. By definition, the expected utility of the agent under the trivial scheme is at least equal to

$$\lambda_1(\bar{a})U(Y_1 - R) + \lambda_2(\bar{a})U(Y_2 - R) - G(\bar{a}).$$

This provides a consumption distribution superior to that under the optimal scheme in the sense of second-order stochastic dominance, because it involves zero cost of audit as well as lower risk. Hence the trivial scheme provides higher expected utility, which is a contradiction.

Next, we prove that $V_2 \geq V_1$. Suppose otherwise: that is, in the optimal scheme, $V_2 < V_1$, and the agent chooses \bar{a} . Using the preceding argument, $T_2 > T_1$, and $p_2 = 0, p_1 > 0$. Consider a variation in which T_1 is increased, and T_2 reduced such that the agent's expected utility from choice of action \bar{a} is unchanged. Let $dV_1 < 0$ and $dV_2 > 0$ be the changes in expected utility corresponding to the two incomes. Since action \bar{a} continues to be available, the expected utility of the agent cannot decrease.

If the agent continues to prefer action \bar{a} , the reduction in risk implies that expected profits of the principal must increase. Further, since $T_2 > T_1$, this variation will preserve the reporting constraints. Thus, by Lemma 1 the original contract could not be optimal.

It remains to consider the case where under the revised scheme the agent prefers some action $\hat{a} \neq \bar{a}$. By definition,

$$\begin{aligned} V_1 + \lambda_2(\bar{a})[V_2 - V_1] - G(\bar{a}) &\geq V_1 + \lambda_2(\hat{a})[V_2 - V_1] - G(\hat{a}); \\ V_1 + dV_1 + \lambda_2(\hat{a})[V_2 - V_1 + dV_2 - dV_1] - G(\hat{a}) \\ &\geq V_1 + dV_1 + \lambda_2(\bar{a})[V_2 - V_1 + dV_2 - dV_1] - G(\bar{a}). \end{aligned}$$

Adding these two inequalities, we obtain $[\lambda_2(\hat{a}) - \lambda_2(\bar{a})][dV_2 - dV_1] \geq 0$, which implies that $\lambda_2(\hat{a}) \geq \lambda_2(\bar{a})$. Since $T_2 > T_1 > T_1 - p_1(E_1 - F_1)$ in the original scheme, this implies that expected profit of the principal must be even higher in the new scheme than in the case where the agent continues to prefer \bar{a} . Since the agent is also better off, we obtain a contradiction.

Since $F_2 = 0, V_2 = U(Y_2 - T_2)$; and since $F_1 \leq 0, V_1 \geq U(Y_1 - T_2)$. Hence $V_2 \geq V_1$ implies that $Y_2 - T_2 \geq Y_1 - T_1$.

Q.E.D.

Do the monotonicity results extend to the case of three or more possible realizations of income? In a setting of pure moral hazard—that is, where the principal may observe the agent's realized income without cost—Grossman and Hart [1983] have shown that consumption in the optimal scheme may be nonmonotone in income. In their setting, it can also be shown that optimal transfers may be nonmonotone. In our setting, if audit costs are sufficiently small, the same results will follow from continuity arguments. Such nonmonotonicities are, however, “driven” by the moral hazard aspect of the problem. So, in the following analysis we focus on a setting without moral hazard—that is, where the probabilities of the various realizations of income are exogenously given— λ_i for income Y_i .

In the pure moral hazard context while Grossman and Hart [1983] have presented examples where the agent's consumption is nonmonotone in his income, they have also shown that it is not possible for consumption to be everywhere decreasing in income. The following Proposition establishes the same result in the case of no moral hazard where auditing is costly. It further shows that the highest income report Y_n is associated with the highest transfer, and is thus never to be audited. This implies that audit probabilities cannot be everywhere increasing, and transfers cannot be everywhere decreasing in income.

PROPOSITION 4. Assume that there is no moral hazard; that is, the probabilities of the various realizations of income are exogenously given. Then in any optimal scheme, the highest income report Y_n is not audited, and is associated with (a) the highest transfer, (b) the highest expected utility, and (c) the highest consumption in the event of no audit.

Proof of Proposition 4. a. We first show that $T_n \geq T_i$ for all i . Suppose otherwise: that the set of incomes $B = \{Y_j | T_j \geq T_i \text{ for all } i\}$, associated with the highest transfer, does not contain Y_n . We claim that in this case we can increase T_n slightly without violating any reporting constraints. Consider any income Y_i for which $T_i < T_n$; clearly, the agent with true income Y_n will only consider reporting an income satisfying this property. Consider any Y_l in the set B paying the highest transfer. Then $T_l > T_n > T_i$. Since U is concave, we have

$$\begin{aligned} U(Y_n - T_n) - U(Y_l - T_l) &> U(Y_n - T_n) - U(Y_l - T_n) \\ &\geq (1 - p_i)[U(Y_n - T_i) - U(Y_l - T_i)]. \end{aligned}$$

As Y_i belongs to B , $p_i = 0$ and hence, by the reporting constraint, $V_i = U(Y_i - T_i) \geq (1 - p_i)U(Y_i - T_i)$. Thus, the agent with true income Y_n must strictly prefer to report truthfully rather than report Y_i . Therefore, we can raise T_n slightly and lower the highest transfer (corresponding to each state in B), without violating any reporting constraints and without reducing expected profits. Since $Y_n - T_n > Y_i - T_i$ for any Y_i in B , this results in reduced risk and thus higher expected utility for the agent, which is a contradiction.

b. Suppose that there exists income Y_j less than Y_n with the property that $V_j > V_n$. Then, for any Y_i , by the reporting constraints, $V_j > V_n \geq (1 - p_i)U(Y_n - T_i) \geq (1 - p_i)U(Y_j - T_i)$; so if true income is Y_j , the agent will strictly prefer to report truthfully. We can therefore increase T_j slightly, without violating any reporting constraints. Simultaneously, we can reduce the transfer associated with type Y_n and also of any other income Y_i with $T_i = T_n$, without violating any reporting constraints (because T_n is the highest transfer) and without reducing expected profits. This reduces risk borne by the agent, and thereby increases expected utility, a contradiction.

c. By Lemma 2, since $V_n \geq V_i$, $U(Y_n - T_n) = V_n \geq V_i \geq U(Y_i - T_i)$, for all i .

Q.E.D.

In order to investigate possible nonmonotonicities of consumption, we have numerically computed the solution to the following problem. Suppose that there are three income levels: $Y_1 = 10$, $Y_2 = 30$, $Y_3 = 40$, with respective exogenous probabilities $\lambda_1 = 0.45$, $\lambda_2 = 0.10$, and $\lambda_3 = 0.45$. Let $U(Y) = Y^{0.4}$, $R = 4$, and $E_1 = E_2 = E_3 = 2$. The computed solution is $T_1 = -10.35$, $T_2 = 9.88$, $T_3 = 17.84$; $p_1 = 0.28$, $p_2 = 0.12$, $p_3 = 0$; $F_1 = -0.60$, $F_2 = -0.68$, $F_3 = 0$. Thus, $Y_1 - T_1 = 20.35$, $Y_2 - T_2 = 20.12$, and $Y_3 - T_3 = 22.16$, and $V_1 = 3.35$, $V_2 = 3.33$, and $V_3 = 3.45$. It therefore appears that consumption in the event of no audit and expected utility may decrease over some range of incomes.⁷

Finally, we should mention that the monotonicity of transfers and audit probabilities in the case without moral hazard, but with three or more possible income levels, remains an open question.

III. CONCLUDING REMARKS

We first summarize the reasons why debt contracts are never optimal in our model. Debt is characterized by a constant transfer

7. Other examples of nonmonotone consumption are available on request from the authors.

paid by the agent, whenever his income is sufficient to afford the payment. In such states, the principal does not expend any resources in verifying the agent's true income. However, the agent declares himself bankrupt whenever his true income falls below the stipulated transfer. In such cases, the principal always investigates the agent's true income, and appropriates it entirely, subject to limits on the agent's liability. Thus, debt contracts are characterized by the following features: (a) there is a cutoff income level above which the principal never audits, and below which he always audits; (b) in the event of not being audited, the transfer from the agent to the principal is independent of his income; (c) in the event of being audited, the agent's consumption is constant; and (d) the agent's consumption and utility level are increasing in his true income level.

Our results imply that optimal contracts may possess none of these features. Property (a) is violated whenever optimal audits are random. The randomness of audits also suggests that the principal should be more active in initiating bankruptcy proceedings (as it is the principal's decision whether or not to audit). In the optimal contract, after observing his realization of income, the agent should prefer the principal to audit. This aspect does not correspond closely to the institution of bankruptcy. Property (b) has been shown to be false in any nontrivial scheme (in the discussion following Proposition 2). Property (c) is clearly invalid in the numerical example studied at the end of the previous section: there the agent obtains consumption levels of 20.95 and 20.80 in income states $Y_1 = 10$ and $Y_2 = 30$, respectively, if he happens to be audited. Finally, the agent's consumption (in the event of no audit) and utility may decrease over some ranges of income, so property (d) does not necessarily hold. While properties (c) and (d) have been shown to be violated only in certain examples, properties (a) and (b) will be violated in general.

Our analysis therefore casts doubt on the model most commonly employed to account for the widespread use of debt. Perhaps more satisfactory explanations might be based on the relative simplicity of debt contracts. We stress, however, that the main positive results of the present paper, namely, that optimal audits are random, and that truthful reports should be rewarded when verified, provide useful guidance in the design of income-tax enforcement. In a related paper [Mookherjee and Png, 1988] we address in detail the problem of designing optimal tax and audit enforcement schemes subject to a requirement of horizontal equity.

The main results of the model are robust with respect to three significant extensions not considered explicitly in this paper. First, the agent may also have private ex ante information about the likelihood of alternative incomes resulting from any given action. The principal may then offer the agent a menu of possible contracts, and induce the latter to self-select into appropriate contracts according to his ex ante information (see Melumad and Reichelstein [1988]). By essentially similar arguments, our results then extend to each possible contract that the principal may offer the agent. Second, we assume throughout that audits are perfect: the principal discovers the agent's true realized income without error. It may be verified, by continuity arguments, that our results extend to the case when there are "small" errors in auditing. Finally, we assume that the principal can always commit to a scheme involving interior audit probabilities, even though in equilibrium, he knows that the agent reports truthfully, and therefore may have an ex post incentive not to carry out the prescribed audit strategy. In a related paper, Melumad and Mookherjee [1988] show that the principal may implement any optimal random audit strategy by delegating responsibility for audits to an independent auditor, and committing instead to a simple incentive contract for this auditor. These contracts induce a game between the agent and the auditor in which the unique sequential equilibrium outcome is the optimal commitment solution.

STANFORD UNIVERSITY
UNIVERSITY OF CALIFORNIA, LOS ANGELES

REFERENCES

- Baiman, S., and J. Demski, "Economically Optimal Performance Evaluation and Control Systems," *Journal of Accounting Research*, XVIII Supplement (1980), 184-220.
- Border, K., and J. Sobel, "A Theory of Auditing and Plunder," *Review of Economic Studies*, LIV (1987), 525-40.
- Dye, R. A., "Optimal Monitoring Policies in Agencies," *Rand Journal of Economics*, XVII (1986), 339-50.
- Gale, D., and M. Hellwig, "Incentive-Compatible Debt Contracts: The One Period Problem," *Review of Economic Studies*, LII (1985), 647-63.
- Grossman, S., and O. Hart, "An Analysis of the Principal Agent Problem," *Econometrica*, LI (1983), 7-45.
- Melumad, N., and D. Mookherjee, "Delegation As Commitment: The Case of Income Tax Audits," *Rand Journal of Economics* (1989), forthcoming.
- , and S. Reichelstein, "Value of Communication in Agencies," Research Paper 895, Graduate School of Business, Stanford University, 1988.
- Mookherjee, D., and I. Png, "Enforcement Costs and Optimal Progressivity of Income Taxes," 1988.
- Prescott, E., and R. Townsend, "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica*, LII (1984), 21-46.

- Reinganum, J., and L. Wilde, "Income Tax Compliance in a Principal Agent Framework," *Journal of Public Economics*, XXVI (1985), 1-18.
- Townsend, R., "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, XXII (1979), 265-93.
- , "Information Constrained Insurance," *Journal of Monetary Economics*, XXI (1988), 411-50.