THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LIV June 2006 No. 2

INCENTIVES FOR CORRUPTIBLE AUDITORS IN THE ABSENCE OF COMMITMENT*

FAHAD KHALIL†
JACQUES LAWARRÉE‡

In the absence of commitment to auditing, we study the optimal auditing contract when collusion between an agent and an auditor is possible. We show that the auditor can be totally useless if the auditor's independence can be compromised with relative ease. Even very stiff sanctions on fraud will be unable to make auditing optimal. We then derive a demand for independent external auditing. We endogenize collusion cost as the cost from the risk of future detection. We also derive a justification for the focus of the recent audit reforms on penalties on CEOs in cases of audit fraud.

I. INTRODUCTION

MORE FREQUENTLY THAN WE WOULD LIKE TO HEAR, the press reports scandals exposing non-diligent auditing. The Enron scandal is simply one in a growing string of embarrassments for the auditing profession. The Security and Exchange Commission (SEC) reports that the number of cases of fraud has increased by 41% between 1998 and 2001.[1] In a recent example, which, prior to the Enron debacle, had been touted as the biggest case of accounting fraud (estimated at $19 billion), the SEC investigated the firm CUC International, the travel and transportation conglomerate that owns the Ramada hotel and Avis car rental chains.[2] It is alleged that CUC fooled Ernst and Young auditors for a number of years and then conspired with them.[3] Such examples typically show that an auditor initially gave a report of compliance for a firm, but subsequent evidence demonstrated wrong-doing by the firm and collusion between the auditor and the firm.

*We thank Patrick Bolton, Mathias Dewatripont, Alfredo Kofman, Roland Strausz, Jean Tirole, Mike Townsend, two anonymous referees and the editor for helpful comments.

†Authors' affiliations: Department of Economics, Box 353330, University of Washington, Seattle, WA 98195, U.S.A.
e-mail: khalil@u.washington.edu
‡Department of Economics, Box 353330, University of Washington, Seattle, WA 98195, U.S.A, and ECARES (Brussels).
e-mail: lawarree@u.washington.edu

[1] Chartier [2002].
[2] *New York Times* [2000].
[3] The CUC International shareholders' complaint states that the accounting firm concealed the fraud, 'first by negligence, then by cautious avoidance and eventually by active facilitation.'

There is a simple solution to the problem of collusion: it is to abandon enforcement altogether. For example, the end of Prohibition in the United States eliminated the need for corrupt payments to the police. Put differently, the solution to a corrupt auditor is to eliminate the *raison d'être* of the auditor. Can this be an optimal response? The current theoretical literature on collusion that followed Tirole [1986] would validate such a solution only when auditing is too costly or too imprecise.[4] In this paper, we present a new rationale: the absence of commitment to auditing. When the principal cannot commit to auditing and the auditor is corruptible, it may not be optimal to audit even if auditing is costless and perfect. This new result stems from the interaction between the commitment and the collusion problems in auditing.

The lack of commitment is a well-recognized issue for auditing. Auditing suffers from a time consistency problem. If the firm being audited never cheats, the audit never reveals any wrongdoing and there are no *ex post* incentives to audit (Khalil [1997]). This problem could be modeled as the auditor's moral hazard. The auditor would not perform a thorough audit if he knew that he could only confirm the agent's report (Baiman, Evans and Noel [1987]). Studying the tax compliance game, Graetz, Reinganum and Wilde [1986] were among the first to model the IRS auditor as a strategic actor who does not commit to an audit policy. Focusing on regulatory and procurement relationships, Laffont and Tirole [1993] argue that either legal prohibitions or the inability to describe future technologies may prevent commitment. Picard [1996] explains that a commitment to audit insurance claims in order to detect fraudulent ones is not easy to achieve, especially when the optimal audit policy is random. In a banking framework, Khalil and Parigi [1998] show that banks may use the loan size to overcome the problem due to a lack of commitment to auditing.

Without commitment, auditing will only take place if it is optimal *ex post*. Cheating must occur in equilibrium so that the principal can expect to collect a penalty to cover auditing costs. That is, in the absence of commitment, it is *as if* the principal had to provide incentive for himself to hire the auditor. The principal must also anticipate that the agent can bribe the auditor. Even if there is a cost of writing an illegal side contract[5] – collusion costs – the auditor may collude with the agent and submit a false report.

To deter collusion, consider a strategy that rewards the auditor for turning down the bribe. This reward must be at least as high as the maximum bribe, which is the penalty net of any collusion costs. If the collusion costs are small, this strategy is very costly for the principal, as he must give up almost the

[4] See for instance Proposition 1 in Kofman-Lawarrée [1993]

[5] This is a standard assumption in the corruption literature. It reflects that illegal collusion contracts are difficult to enforce and therefore costly to implement. In the second part of the paper, we endogenize this cost.

entire penalty as a reward to the auditor. At the same time, the principal still faces his own incentive problem, i.e., inducing himself to hire the auditor. In equilibrium, the principal must induce shirking with high probability in order to make up for the cost of an audit with the expected penalty. But then, the productivity of the agent is lost (almost entirely) and the principal is better off with the second-best contract without auditing. We show that given any penalty, if collusion costs are sufficiently small, auditing is never optimal. This is fairly surprising, for a rather robust result from the literature is that auditing is always optimal if the penalty is high enough whether collusion is possible or not.[6]

Suppose now that the principal can also rely on another signal to deter collusion. For example, an economic downturn exposes an unprofitable firm that had nonetheless received strong marks from the auditor; the contractor and building inspector's fraud is revealed by the collapse of a building in a mild earthquake; an unrelated audit by the IRS may expose fraudulent labor practices. A more interesting case, considered in this paper, is when the principal is allowed to buy such external signal, which we model as an external auditor's report. The external auditor is honest but more expensive than our first auditor – now called internal auditor. We show that for large enough penalties, shirking, collusion and detection occur in equilibrium as seen in press reports such as the one reported earlier in this paper. We also find that when the principal uses an external auditor, he no longer rewards the internal auditor.

Thus, we derive collusion cost endogenously by interpreting collusion cost as the cost stemming from the risk of future detection. Only recently have some authors endogenously derived collusion costs, but these contributions do not rely on the threat of future detection. Using a dynamic model with reputation, Martimort [2000] endogenously derives the cost of writing side contracts. Faure-Grimaud, Laffont and Martimort [1999, 2002] show that the cost of collusion between supervisor and agent depends upon the collusion stake, the accuracy of the supervision technology and the supervisor's degree of risk aversion.

The recent accounting scandals have led to demands for reforms of the auditing industry. The Sarbanes-Oxley Act of 2002[7] – also known as the corporate corruption bill – is a main component of recent reforms. There is also a debate in the auditing profession about increasing the liability of auditors (Grout *et al.* [1994]). Our contribution to the debate is to point out that the penalty on the auditors could be thought of as a penalty on collusion and the penalty on the agent as a penalty on non-compliance. Since the central incentive problem is shirking, and not collusion, increasing the penalty on the agent is more effective in decreasing shirking and improving

[6] See Baron and Besanko [1984] and Kofman-Lawarrée [1993].
[7] U.S Congress [2002].

welfare than increasing the penalty on the auditor. Therefore, our results may provide a theoretical basis for the strong emphasis on penalties on CEOs and CFOs in the Sarbanes-Oxley Act.

Related Literature

Very few authors have looked at the dual problem of commitment and collusion in auditing. An exception is Strausz [1997a]. While the model and focus of the two papers are different, he also finds that collusion may be optimal[8] if the principal cannot commit not to renegotiate. In his model, an agent works on a project whose cost is known only to the agent and a supervisor (internal auditor in our model), but there is no explicit productive action. The supervisor provides a report of cost and, in addition, the principal receives a free but imperfect signal of collusion, which can be interpreted as the report of a free external auditor. The penalty in his model is the shutdown of the project, which is costly for the principal and the source of the problem of credibility. By inducing collusion, the supervisor's report is rendered valueless and this makes it possible for the principal to shut down the project based *only* on the (external) signal while keeping the contract renegotiation-proof. Thus, although it is a very rich model, it does not allow him to characterize the optimal amount of shirking nor the equilibrium without an external auditor. More importantly, when collusion occurs in our model, both the internal and the external auditors' reports are used in the contract. Since in our model, the problem of credibility is due to the fact that audits are costly, our focus is more on how to make audits credible and the implications on optimal contracts. Also, with exogenous penalties, we are able to investigate the types of penalties that are more effective.

In contrast to the traditional timing of auditing models, Lambert-Mogiljansky [1994] introduces a monitor *before* the productive action takes place. In her model, not only can the principal not commit to audit probabilities, but he cannot even commit to the terms of the contract. Her results also emphasize the importance of rent dissipation. However, unlike our paper, she finds that it may be optimal to allow collusion in equilibrium with only one internal auditor.[9]

Recent papers have looked at the introduction of multiple auditors to control collusion (Kofman and Lawarrée [1993, 1996] and Laffont and Martimort [1999]). Once again, they all assume commitment to auditing. Kofman and Lawarrée [1996] and Laffont and Martimort [1999] introduce a second *internal* auditor. Kofman and Lawarrée [1993] are closer to our

[8] Other papers in which collusion may be optimal, but for different reasons, are cited in this paper.

[9] Note that we ignore the issue of delegation and the effect of collusion. On this topic, see Strausz [1997b].

model as they investigate the introduction of an *external* auditor. They show that if auditing were without error, no external auditor would be hired. Our contribution here is to derive a demand for external auditors that is sequentially optimal even with error-free audits.

The paper is organized as follows. In Section 2 we present the basic model of auditing without commitment. In Section 3, we introduce the possibility of collusion and derive the optimal contract. In Section 4, we introduce an external auditor. We conclude in Section 5.

II. THE BASIC MODEL

We present a model of adverse selection with effort and monitoring. A risk neutral principal hires a risk neutral agent. The agent can have low productivity θ_1, or high productivity θ_2, with $\theta_2 > \theta_1 > 0$. The agent puts in non-negative effort e which, together with his productivity parameter, determines profit or output $X = \alpha(\theta, e)$, where $\alpha_e > 0$, $\alpha_{ee} < 0$, $\alpha(\theta_2, e) > \alpha(\theta_1, e)$, and $\alpha_e(\theta_2, e) > \alpha_e(\theta_1, e) > 0$. The principal collects output and pays a transfer t to the agent. The cost to the agent of exerting an effort e is $\psi(e)$, where $\psi_e > 0$, and $\psi_{ee} > 0$. To obtain strictly positive but bounded optimal efforts, we also assume $\alpha(\theta, 0) = 0$, $\psi(0) = 0$, $lim_{e \to 0}\, \psi_e(e) = 0$, $lim_{e \to 0}\, \alpha_e(\theta, e) = \infty$, $lim_{e \to \infty}\, \psi_e(e) = \infty$ and $lim_{e \to \infty}\, \alpha_e(\theta, e) = 0$. The agent receives t_i for producing $X_i = \alpha(\theta_i, e_i)$ for $i = 1, 2$. The agent's reservations utility is assumed to be zero. If there was full information, production would be efficient, and there would be no rent:

$$\alpha_e(\theta_i, e_i^*) = \psi_e(e_i^*),$$
$$t_i^* = \psi(e_i^*).$$

We will also refer to the above as the first-best efforts and transfers.

Under asymmetric information, the contract specifies the outputs and transfers for each state. It is common knowledge that the principal assigns the probability q to the event that a particular agent is of type θ_1. The agent knows θ before he signs the contract and chooses his effort. While the output is publicly observable, both e and θ are the private information of the agent. This gives an opportunity for the high type to shirk. Shirking means that the high type can mimic the low type by producing the output designated for the low type, $\alpha(\theta_1, e_1)$. If the high type shirks, he must put in effort $\hat{e}_1$, where $\alpha(\theta_2, \hat{e}_1) \equiv \alpha(\theta_1, e_1)$.

After the output is produced and publicly observed, the principal can order an audit at a cost z to find out if the agent shirked. Therefore, we are assuming that the principal cannot commit to auditing before the output has been revealed[10]. In this framework an audit will either reveal e or θ without error.

[10] We assume that the principal cannot get around the commitment problem by imposing a large penalty on himself. U.S. courts do not enforce penalties designed to spur actions.

If there is an audit, and it reveals shirking, the agent is assessed an exogenous penalty P.[11] It can be verified that, in equilibrium, a low-type agent does not shirk and the principal does not audit when output is high. The principal is only concerned about a high type agent shirking. We allow for mixed strategies: the probability that a high-type agent will shirk is denoted by m, and the probability that the principal will audit after X_1 is produced is denoted by γ. Upon observing a low output, the principal can compute the probability that a high-type agent has shirked. We denote this probability ϕ, with

$$\phi = (1-q)m/[q + (1-q)m].$$

From now on, by *random shirking* we mean that the high type produces X_1 with probability m and X_2 with probability $(1\text{-}m)$.

We summarize the above with the timing:

1. Nature chooses θ, and only the agent learns it.
2. The principal offers the agent a contract.
3. The contract is accepted if it guarantees the agent his opportunity profits (normalized to zero) in each state.
4. The agent chooses effort and output is produced.
5. The principal collects output and pays the agent the transfer.
6. The principal decides whether to audit.

If the principal could commit to auditing probabilities, the solution to this problem is well known (see e.g., Baron-Besanko [1984]). Since he cannot commit, the contract has to give him incentive to perform the audit *ex-post*, which is possible by inducing the agent to shirk in equilibrium.[12]

The principal chooses the contract $\{e_1, e_2, t_1, t_2\}$. Bester and Strausz [2001], in an insightful paper, show that with a single agent, optimal mechanisms can always be represented by direct mechanisms even in the absence of commitment. There is no such result available for the case of multiple agents (Bester and Strausz [2000]). Therefore, in subsequent sections with a strategic auditor, we derive the optimal contract among the class of contracts where the cardinality of the messages for the agent is equal to the cardinality of the type space.

When writing the principal's problem it is convenient to let the principal choose $\{e_1, e_2, t_1, t_2, m, \gamma\}$, but making sure that m and γ are sequentially

[11] In this setup, the principle of maximum deterrence applies. Therefore, without an exogenous upper bound, the first-best can be approximated with very large penalties. See, e.g., Baron-Besanko [1984].

[12] A more elaborate discussion of the issues in this section can be found in Khalil [1997]. However, the main sections of that paper are presented using *transfer dependent* penalties. Since our focus here is on the dual problem of commitment and collusion, we simplify the exposition by modeling penalties as independent of the transfer. With transfer dependent penalties, we would find over-production by the low-type.

optimal. The principal's problem P_N is stated next.

$$\begin{aligned}&\text{Max}[q + (1-q)m][\alpha(\theta_1, e_1) - t_1 + \gamma(\phi P - z)]\\&\quad + (1-q)(1-m)[\alpha(\theta_2, e_2) - t_2]\end{aligned}$$

s.t.

$$(2\text{IR}_1) \quad t_1 - \psi(e_1) \geqslant 0,$$

$$(2\text{IR}_2) \quad (1-m)[t_2 - \psi(e_2)] + m[t_1 - \psi(\hat{e}_1) - \gamma P] \geqslant 0,$$

$$(2\text{IC}_m) \quad m \in \text{argmax}_{\text{m}'}(1 - \text{m}')[\text{t}_2 - \psi(\text{e}_2)] + \text{m}'[\text{t}_1 - \psi(\hat{\text{e}}_1) - \gamma\text{P}],$$

$$(2\text{IC}_\gamma) \quad \gamma \in \text{argmax}_{\gamma'}\, \gamma'[\phi\text{P} - \text{z}].$$

The objective function is the principal's payoff. The probability that X_1 will be produced is $[q+(1-q)m]$, while X_2 is produced with probability $(1-q)(1-m)$. If X_1 is produced, the principal audits with probability γ, and then he collects the penalty with probability ϕ, where the cost of an audit is z. The constraints (2IR_1) and (2IR_2) are the two individual rationality constraints. Constraint (2IR_2) takes into account that the high-type agent may shirk with probability m. The two incentive compatibility constraints guarantee sequential rationality. The constraint (2IC_m) ensures that m maximizes the high-type agent's payoff given the contract. The constraint (2IC_γ) ensures that γ is optimal for the principal after X_1 is produced.

In the remainder of this section our main objective is to establish properties of optimal contracts that induce random audit, and show that random audit is optimal when the penalty is high enough. However, depending on parameter values, the solution to problem P_N may or may not involve random audit.

Proposition 0. (a) If audits are not optimal ($\gamma = 0$), the solution to P_n is the second-best contract, characterized as follows:

$$\begin{aligned}&\psi_e(e_2) = \alpha_e(\theta_2, e_2),\\&\psi_e(e_1) = \alpha_e(\theta_1, e_1) - [(1-q)/q][\psi_e(e_1) - \psi_e(\hat{e}_1)(\alpha_e(\theta_1, e_1)/\alpha_e(\theta_2, \hat{e}_1)],\\&t_1 = \psi(e_1), t_2 = \psi(e_2) + \psi(e_1) - \psi(\hat{e}1), m = 0,\end{aligned}$$

and the principal's payoff is strictly smaller than under full information or first best.

(b) Certain audit ($\gamma = 1$) is optimal if and only if shirking occurs with certainty ($m = 1$).

(c) If random audits ($0 < \gamma < 1$) are optimal, there is random shirking by the high type ($0 < m < 1$), but the contracted quantities are efficient. There is no rent for either type.

(d) There exists a penalty level L, such that for $P > L$ random audits $0 < \gamma < 1$ are optimal.

Proof. In appendix A.

If there is no audit, as described in part (a), the revelation principle applies since our 'no commitment' assumption only pertains to auditing. The revelation principle implies that the principal can do no better than deter shirking. As is well known, this is best accomplished by offering the second-best contract.[13] In the second-best contract, the high type produces efficiently but receives rent and the low type under-produces and receives no rent.

On the other hand, as described in part (b), we may have a case when the high type shirks with certainty even though an audit is certain to follow. Clearly this can happen only if the penalty is smaller than the rent from shirking (gross of the penalty). Since optimal efforts and transfers are bounded, this case cannot be optimal for high enough penalties.

When a random audit is optimal, as described in part (c), the agent will not want to shirk with certainty; otherwise the principal would want to audit with certainty and since the penalty is high, the agent would not shirk. Also, unless the agent is shirking with some probability, the principal will not audit. Thus, there must be random shirking under random auditing.

In this mixed strategy equilibrium, the principal is indifferent between auditing and not auditing, i.e., given X_1 the net expected return from an audit is zero. The high-type agent also is indifferent between shirking and not shirking. Since the expected return from an audit is zero, an increase in the probability of audit does not directly affect the principal's payoff, but indirectly it reduces rent through the ($2IC_m$). Thus, the probability of audit is increased until rent is reduced to zero, and there is no rent in either state.

Since there is no rent, the usual rent versus efficiency trade-off is no longer present. The probability of shirking is determined by ($2IC_\gamma$), where ϕP–$z = 0$. Since m is a function of parameters alone, and there is no rent, there is no reason to distort efforts from the first best. Note that, unlike models with commitment, observed output always corresponds to the efficient output level for one state or the other.

Finally, in part (d) of proposition 0, we show that random audit is optimal if the penalty is high enough. The principal's payoff under random audit is different from the first-best payoff only due to shirking. As the penalty

[13] See for example, Laffont and Tirole [1993].

becomes very large, the probability of shirking goes to zero. Also, as mentioned above, certain audits cannot be optimal for high penalties.

III. CORRUPTIBLE AUDITOR

We now consider possible collusion between the auditor and the agent. Collusion means that the auditor does not report detection of shirking presumably in exchange for a bribe from the agent.

The only way the principal can prevent the auditor from accepting a bribe is to make sure that it is not profitable for the auditor to do so. There are essentially two schemes available to the principal for this purpose. The first scheme would threaten the auditor with a strong punishment if he were caught accepting a bribe. However, in order to discover that bribing occurred, the principal needs a signal or another auditor to control the first one. We examine such a scheme in section 4. Another possible way to prevent bribing is to reward the auditor every time he turns down a bribe. In this section, we let the principal use rewards to deter collusion. The principal pays the auditor $(z + w)$ when shirking is reported, and only z otherwise. The reward w is assumed to be non-negative.

The principal offers the contract $\{e_1, e_2, t_1, t_2, w\}$. Given the contract, the agent decides on his strategy of shirking. Once the output is revealed, the principal decides whether to use the auditor, and then the auditor and agent decide upon the bribe, denoted by B.[14]

We assume that each time the auditor has incriminating evidence that would convict the agent, the agent knows it.[15] We give total bargaining power to the auditor when the auditor and agent negotiate their side contract.[16] Since the agent pays a penalty P when shirking is detected and reported, the bribe will be the maximum he is willing to pay, i.e., it will be equal to P. We want to account for collusion costs and, as in Laffont-Tirole [1993, chapter 11], we assume that a bribe of \$1 is worth only \$$\lambda$ to the auditor, with $0 \leqslant \lambda \leqslant 1$. The cost of collusion may arise due to the risk of legal sanctions or due to the cost of writing and enforcing an illegal side contract.[17] Since the collusion cost is proportional to the bribe, $(1–\lambda)$ measures the cost of collusion per unit of bribe in this model.

[14] Kessler [2000] shows that the timing of the game matters for the relevance of collusion.

[15] This is direct consequence of the assumption of perfect auditing. It also simplifies the bargaining problem between the auditor and the agent. They both bargain over the amount of the bribe under symmetric information. This assumption is, however, realistic in an auditing framework as the agent can observe the auditor's searching process and therefore deduce his conclusions.

[16] This assumption allows us to examine the case where collusion is more costly to prevent, which is an adequate benchmark.

[17] Later we discuss the cost of collusion as stemming from the future threat of detection.

The principal's problem with collusion — labeled P_{NC}— is presented below.

$$\begin{aligned}&\mathrm{Max}\{[q+(1-q)m][\alpha(\theta_1,e_1)-t_1+\gamma[\phi(P-w)-z]\}\\&\quad+(1-q)(1-m)[\alpha(\theta_2,e_2)-t_2]\end{aligned}$$

s.t.

$(3\mathrm{IR}_1)$ $\quad t_1-\psi(e_1)\geqslant 0,$

$(3\mathrm{IR}_2)$ $\quad (1-m)[t_2-\psi(e_2)]+m[t_1-\psi(\hat{e}_1)-\gamma B]\geqslant 0,$

$(3\mathrm{IC}_m)$ $\quad m\in \mathrm{argmax}_{\mathrm{m}'}(1-\mathrm{m}')[\mathrm{t}_2-\psi(\mathrm{e}_2)]+\mathrm{m}'[\mathrm{t}_1-\psi(\hat{\mathrm{e}}_1)-\gamma\mathrm{B}],$

$(3\mathrm{IC}_\gamma)$ $\quad \gamma\in \mathrm{argmax}_{\gamma'}\ \gamma'[\phi(\mathrm{P}-\mathrm{w})-\mathrm{z}].$

(CIC) $\quad w\geqslant\lambda B$

(B) $\quad B=P$

The incentive constraint $(3\mathrm{IC}_m)$ for the agent and the IR constraints remain the same as in section 2 since, instead of the penalty, the agent pays P as a bribe if shirking is detected. We argue next that collusion will be deterred, and this is reflected in the new coalition incentive compatibility constraint (CIC) in P_{NC}. If an audit occurs, the principal now receives $(P-w)$ if shirking is reported and nothing otherwise. If $w<\lambda P$, collusion will occur and shirking will not be reported. Since an audit costs z, the principal will only use an auditor if he can prevent collusion; otherwise an audit is not *ex-post* optimal. Therefore, we have the following Lemma.

Lemma 1. If it is optimal to audit, then it is optimal to deter collusion, i.e., the principal will set $w\geqslant\lambda P$.

This result is in contrast with the result under commitment to auditing. With commitment, the principal can use the auditor even if it is not *ex-post* optimal. Then the principal deters shirking, and therefore, never receives the penalty.[18] The threat of penalty is used to lower rent, and it does not matter if the threat is a penalty or if it is a bribe. They are equally effective in lowering rent. The principal is indifferent between deterring and allowing collusion in that case.

Comparing with our maximization problem without collusion, two new features appear. First, the constraint (CIC) ensures that the optimal contract will deter collusion. Second, when the principal detects shirking, he now

[18] The principal may receive some penalty if there is error in auditing, but all parties anticipate this and the principal has to compensate for these errors *ex-ante*.

collects $P - w$ instead of P. These changes, however, do not affect the proof of Proposition 0 (parts a and b) which remains valid. Results similar to part (c) also hold, but we have to account for the possibility of collusion:

Proposition 1. If random audits are optimal, there is random shirking by the high type, but the contracted quantities are efficient. There is no rent for either type. The threat of collusion does not affect the probability of audit, but it increases the probability of shirking, and it lowers the principal's payoff.

Proof. In Appendix B.

The second stage equilibrium differs when collusion is possible compared to when it is not. The principal must pay the extra amount w to the auditor to deter collusion and the *ex-post* returns of detecting shirking goes down. The principal is indifferent between auditing or not, and this indifference condition is now given by

$$\phi(P - w) = z,$$

instead of $\phi P = z$ in section 2 where collusion was not possible. Remembering that $\phi = (1 - q)m/[q + (1 - q)m]$, we know that there will be more shirking (m is higher) when collusion needs to be deterred ($w > 0$). That is, the agent shirks with a higher probability to keep the principal indifferent between auditing and not auditing.

Note that by lowering w, the principal can induce a reduction in the probability of shirking m. Therefore, (CIC) is binding and $w = \lambda P$, which implies that the principal's indifference condition is now $\phi(1 - \lambda)P = z$, and we have

$$m = qz/[(1 - q)(P - \lambda P - z)].$$

We again see that m is determined by exogenous variables, and therefore, just as in section 2, efforts are efficient when there is no shirking and there is no rent in either state. Then, the constraint ($3IC_m$) implies that the probability of auditing is not affected by the possibility of collusion. The higher m implies that the principal's payoff is lowered.

Proposition 2: When the principal cannot commit to auditing, the threat of collusion can make auditing sub-optimal if λ is high enough, even when large punishments are available, i.e., for any P, there exists $\lambda(P) \in (0, 1)$ such that $0 < \gamma < 1$ is optimal if $\lambda < \lambda(P)$, while $\gamma = 0$ is optimal if $\lambda > \lambda(P)$.

Proof. In Appendix B.

Since $w = \lambda P$, the principal collects $P – w = (1 - \lambda)P$ when shirking is detected. Thus, the probability of shirking must increase with λ to maintain expected (collected) penalty equal to audit cost. As λ increases and raises shirking, auditing becomes less and less attractive and, at some point, the second best contract yields a higher payoff. Note that it is the joint action of the lack of commitment and the possibility of collusion that implies this result. Without the possibility of collusion, as in section 2, high penalties are sufficient to make auditing optimal. Again, we emphasize that this is a new and fairly surprising result, for a rather robust finding from the literature is that auditing is always optimal if the penalty is high enough whether collusion is possible or not.

A more striking feature of the problem of collusion and no-commitment is best shown under the assumption of $\lambda = 1$. Here, the principal does not collect anything after shirking is detected. Since he will never recover the audit cost z, the following result follows.

Corollary 1: If $\lambda = 1$, auditing is not feasible (credible).

The problem arises because the principal has only one instrument to solve two incentive problems. The instrument is the penalty collected from the shirking agent. With this penalty, the principal has to prevent the auditor from colluding and give *himself* incentive to use an auditor. When $\lambda = 1$, the auditor must be given the whole penalty to refuse the agent's bribe and the principal has no (*ex post*) incentive to audit because he cannot recover the cost z.[19]

IV. EXTERNAL AUDITOR

In line with much of the literature, we have so far assumed an exogenous cost of collusion. Informal arguments to motivate this cost typically rely on the risk of future detection of the collusion. In this section we take a first step in deriving collusion cost by explicitly modeling the possibility of future detection. We introduce an external signal that reveals the type of the agent and therefore reveals any shirking or collusion that may have occurred.

Suppose that the principal obtains the external signal by sending another auditor who would police the first one. We call this new auditor 'external' to contrast him with our original auditor whom we now call 'internal'. By assumption, the external auditor is honest, i.e., whatever bribe he is offered, he will refuse it and report the truth to the principal. This honesty comes at an extra cost to the principal.[20] We assume that the external auditor costs z_e (with $z_e > z_i$, the cost of the internal auditor). If the external auditor detects

[19] A scheme paying the auditor P when he catches the agent can only be optimal if $z = 0$.

[20] Using the internal auditor has a lower opportunity cost since he also fulfils multiple functions in the firm, such as filing taxes.

collusion, the agent pays a penalty P^a and the internal auditor pays a penalty P^i.

The benefit of external audits is that they allow the principal to alleviate the problem due to lack of commitment. We will also show that collusion occurs in equilibrium since the principal cannot commit to sending an external auditor. We find that the penalty on the agent can be thought of as a penalty for shirking and the penalty on the internal auditor as a penalty on collusion. We argue that the penalty on the agent is the more effective tool as it attacks the root problem, which is shirking. This could be a justification for the emphasis that recent audit-industry reforms have put on CEO penalties in case of accounting fraud.

In order to highlight the role of the external auditor, we assume that there is no exogenous collusion cost. Indeed, corollary 1 then implies that the internal auditor by himself is useless for the principal. We assume that the principal cannot commit to sending any auditor (internal or external). Once output is realized, the internal auditor is used with probability denoted by γ. The probability that the internal auditor and the agent collude is given by μ. If the internal auditor gives a no-shirking report, the principal does not know if the auditor is telling the truth or if he has been bribed. His posterior belief about the latter event is given by

$$\phi^{\mu} \equiv (1-q)m\mu/[(1-q)m\mu + q].$$

The principal then sends the external auditor with probability denoted by δ. At the collusion stage, the internal auditor, having the entire bargaining power, is able to extract the highest possible bribe B from the shirking agent. Therefore the maximum bribe, B, is such that:

$$B + \delta P^a = P^a \text{ or } B = (1-\delta)P^a,$$

i.e., the maximum bribe equals the expected exoneration from the punishment.

The maximization problem can now be written as

$$\begin{aligned} \text{Max } (q+(1-q)m)\{&\alpha(\theta_1, e_1) - t_1 + \gamma[\phi^{\mu}[(1-\mu)(P^a - w) \\ &+ \mu\delta(P^a + P^i - z_e)] - (1-\phi^{\mu}) - z_i]\} \\ &+ (1-q)(1-m)[\alpha(\theta_2, e_2) - t_2] \end{aligned}$$

s.t.

$$(\text{EA-IR}_1) \quad t_1 - \psi(e_1) \geqslant 0,$$

$$\begin{aligned} (\text{EA-IR}_2) \quad &(1-m)[t_2 - \psi(e_2)] + m[t_1 - \psi(\hat{e}_1) \\ &- \{\gamma(1-\mu)P^a + \mu[B + \delta P^a]\} \geqslant 0, \end{aligned}$$

$$(\text{EA-B}) \qquad B = (1-\delta)P^a,$$

$$(\text{EA-IC}_m) \qquad m \in \operatorname{argmax}_{m'}(1-m')[t_2 - \psi(e_2)] + m'[t_1 - \psi(\hat{e}_1) - \gamma\{(1-\mu)P^a + \mu[B + \delta P^a]\}$$

$$(\text{EA-IC}_\gamma) \qquad \gamma \in \operatorname{argmax}_{\gamma'}\gamma'[\phi^\mu[(1-\mu)(P^a - w) + \mu\delta(P^a + P^i - z_e)] - (1-\phi^\mu)P^i - z_i]$$

$$(\text{EA-IC}_\mu) \qquad \mu \in \operatorname{argmax}_{\mu'}\mu'[B - \delta P^i] + (1-\mu')w$$

$$(\text{EA-IC}_\delta) \qquad \delta \in \operatorname{argmax}_{\delta'}\delta'[\phi^\mu(P^a + P^i) - z_e]$$

We summarize the solution in the following Proposition and leave the proof to the appendix:

Proposition 3. If P^a and P^i are large enough, and z_i is small enough, the optimal contract involves random internal and external audits, random shirking by the high type and random collusion. Again the contracted quantities are efficient. There is no rent for either type and the internal does not receive a bonus ($w = 0$).

Proof. In Appendix C.

We find that it is optimal to use both the internal and external auditor under the familiar condition about large punishments. This result is new and is in contrast to what happens in a model with commitment to auditing. Kofman and Lawarrée [1993] show that in a similar setting an external auditor will not be used if auditing is perfect and the principal can commit to auditing. Then the only role of an external auditor is to reduce the cost of auditing errors to the principal.

The new condition that z_i is small is made precise in the proof. Intuitively, it ensures that the principal will not find it more profitable to hire the external auditor alone, bypassing the internal auditor.[21]

Since the reward (w) of the internal auditor is zero, we find that the principal prefers to punish the internal auditor when convicted of collusion rather than rewarding him for turning down a bribe. Indeed, there are two ways to obtain truthful revelation from the internal auditor. The principal can pay a reward (w) that costs $w = P^a$; or he can use the external auditor with probability δ for a total cost of δz_e. If $w = P^a$, collusion is deterred (i.e.,

[21] Note that $z_i < z_e$ is not sufficient to guarantee the use of an internal auditor since the external auditor is honest.

$\mu = 0$), but then, by (EA-IC$_\delta$)δ would equal zero as the use of an external auditor is not credible. Without collusion cost, corollary 1 implies no audits in this case. We know then that $w < P^a$ if the external auditor is to be used.

We now explain why $w = 0$ at the optimum. In the mixed strategy equilibrium, the internal auditor is indifferent between accepting a bribe or reporting shirking and receiving the reward w. If w is increased, the bribe net of expected penalty from being detected by the external auditor must increase to maintain the internal auditor's indifference. This implies that the external auditor must be used less often (δ falls), which in turn implies that the agent will shirk more.

The driving force behind Proposition 3, which reverses the result of corollary 1, is that the principal has the use of new instrument: the punishment on the internal auditor, P^i. There is another notable difference in the findings compared to section 3, where collusion cost was assumed exogenous. Here, collusion is allowed while in Proposition 2, it is not. The principal has to allow collusion if he is to credibly threaten to use the external auditor. With exogenous collusion costs, as in section 3, this second commitment problem associated with detecting collusion is not present and collusion can be deterred.

We can now reinterpret the collusion cost in terms of penalties brought up by external auditing. By comparing the probabilities of shirking with and without an external auditor we obtain an interpretation of the collusion cost $(1 - \lambda)$ in terms of audit costs and penalties,

$$1 - \lambda = \frac{z_i[(z_e + z_i)(1 - q) + q(P^i + P^a)]}{q(P^i - z_e)z_i + P^a(z_e + z_i)}.$$

Collusion costs could also be derived in a model where a random exogenous signal (such as a mild earthquake or an economic downturn as mentioned in the introduction) reveals the occurrence of collusion. If the principal can only rely on such an exogenous signal, collusion would be deterred in equilibrium since there is no commitment issue with an exogenous signal.

There is a large body of literature that shows that it can be costly to raise the maximum penalty, or why infinite penalties are not optimal.[22] Our model allows us to answer an interesting question: which penalty (P^i or P^a) is the more effective instrument for the principal? What we show next is that the marginal benefit to the principal of a higher P^a is greater than the marginal benefit of a higher P^i.

Proposition 4. An increase in the penalty for the briber is more effective in increasing welfare compared to an increase in the penalty for the bribe receiver.

[22] Garoupa [1997], Shavell [2004].

Proof: In Appendix C.

In the appendix, we show that the principal's payoff is increased more by a marginal increase in P^a than by a marginal increase in P^i. At the limit, we also show that the probability of shirking approaches zero and the principal's payoff becomes arbitrarily close to the first best when P^a tends to infinity. On the other hand, the principal cannot achieve this result when P^i tends to infinity. The result hinges on the relative effects of the two penalties on the probability of shirking, which is the crucial incentive problem. As P^i tends to infinity, we show that m tends to $qz_i/(1-q)(P^a - z_i)$, which is identical to the probability of shirking when collusion is not allowed (section 2).[23] Intuitively, one can think of P^a as a punishment for shirking and bribing, while P^i is a punishment for bribing only. A large P^i will reduce the probability of bribing (μ) arbitrarily close to zero because the internal auditor fears the large penalty.

Our model may provide a theoretical foundation for the recent sweeping changes in the U.S. financial reporting laws. In 2002, the U.S. Congress passed the Sarbanes-Oxley Act to reform the accounting industry as a response to many accounting scandals such as the Enron/Arthur Andersen disaster. The Sarbanes-Oxley Act created an independent auditing oversight board to monitor the auditors. However, one of the most significant provisions of the Act was to increase the penalties for the CEOs and CFOs who falsely certify financial statements of their firm. Indeed, the Act now requires the CEOs and CFOs of the largest companies to personally approve their company reports, which include financial statements. For a 'knowing' false certification, officers now face penalties of one million dollars and/or up to ten years imprisonment, and they face five million dollars and/or twenty years imprisonment for a 'willful' violation. Assuming that accounting fraud occurs to hide shirking or non-compliance and that Congress wants to maximize shareholder value by limiting such non-compliance and fraud, our analysis suggests that Congress was correct in emphasizing penalties on CEOs.

V. CONCLUSION

Two of the more significant issues facing the audit industry are auditor independence and auditor liability. Our paper sheds light on both these issues. We model the interaction of two important problems in providing incentives in auditing: the lack of commitment to auditing and the possibility of collusion between the auditor and the agent. The current literature has by and large ignored the simultaneous presence of these two problems, which, we show, has significant effects on the optimal contract.

One of our main results is the non-optimality of audits under lack of commitment and collusion. It says that if auditor independence can be

[23] Making P^i infinitely large is equivalent to making the external auditor free.

compromised with relative ease, even very stiff sanctions on fraud will be unable to make the use of auditing optimal. Our analysis suggests the importance of increasing collusion cost to improve audit efficiency.

The literature has motivated collusion cost from various sources. In our model, we endogenize collusion cost and interpret it as the risk of being detected by an external signal. For example, the Board of Directors can also order audit by reputable external audit firms because they are not happy with firm performance or suspect fraud. We therefore derive a demand for external audits by reputable firms.

We show that the penalty on the agent can be thought of as a penalty for fraud, while the penalty on the colluding auditor can be interpreted as a penalty for collusion, and that the penalty on the agent is more effective in increasing overall efficiency. This is consistent with the emphasis put on CEO penalties in the audit-industry reforms (Sarbanes-Oxley Act 2002) following the recent accounting scandals.

Our analysis suggests that we may also expect the principal to try to increase his commitment power. One substitute for the principal's commitment ability is for an industry to impose mandatory audits on itself, which may be one explanation for the presence of mandatory audits in the financial world.

An attractive feature of our model is the presence of collusion in equilibrium.[24] Therefore, the presence of collusion in equilibrium is not necessarily a sign of inefficiency of an organization. A technical characteristic of models with commitment is that, because of the revelation principle, the solution exhibits no shirking or collusion in equilibrium. Under no commitment, the revelation principle cannot be used, and, much as in headlines or news reports, shirking, collusion and conviction emerge in equilibrium.

Finally, we have assumed that external auditors are honest, which can be called into question in the light of recent events. On the other hand, the seriousness with which restoring credibility of external audits is being pursued indicates that our assumption may not be entirely misplaced. Thus our assumption boils down to assuming the existence of institutions that bestow a big reward on auditors that can create a reputation of diligence and independence.

APPENDIX A

Proof of Proposition 0

Part (a). If $\gamma = 0$, the revelation principle applies and it implies that ($2IC_m$) will induce $m = 0$. The rest of the proof is standard, and therefore omitted.

[24] Che [1995], Itoh [1993] and Kofman-Lawarrée [1996] are early papers where collusion occurs in equilibrium.

Part (b). (i) $\gamma = 1 \Rightarrow m = 1$. Suppose to the contrary that $\gamma = 1$, but $m < 1$. We can show that $\gamma = 1$ and $m < 1$ cannot be an equilibrium. The ($2IC_\gamma$) constraint implies that $\gamma > 0$ only if $m > 0$. But then $0 < m < 1$, and ($2IC_m$) implies that the high-type agent is indifferent between shirking or not shirking. The principal can break this tie and increase his payoff discontinuously by a small change in the contract inducing the agent to choose $m = 0$. But then γ must be zero.

(ii) $m = 1 \Rightarrow \gamma = 1$. If $m = 1$, $\phi = 1 - \mathrm{q}$. Then $2IC\gamma$ implies that $\gamma = 1$ if $(1 - \mathrm{q})P \geqslant z$, and $\gamma = 0$ otherwise. But by part (a), $\gamma = 0 \Rightarrow m = 0$, therefore $m = 1 \Rightarrow \gamma = 1$.

Part (c). Since $\gamma > 0$, ($2IC_\gamma$) $\Rightarrow m > 0$. Part b then implies $m < 1$. Therefore, the constraints ($2IC_\gamma$) and ($2IC_m$) become indifference conditions associated with a mixed strategy equilibrium. Remembering that $\phi = (1 - q)m/[q + (1 - q)m]$, the two indifference conditions become:

$$(2IC'_m) \quad t_2 - \psi(e_2) = t_1 - \psi(\hat{e}_1) - \gamma P,$$

$$(2IC'_\gamma) \quad P(1 - q)m/[q + (1 - q)m] = z,$$

which define γ and m in terms of the efforts and transfers. Replacing γ in the objective function, and γ and m in the (2IR) constraints, and solving for m to obtain ($2IC''_\gamma$), the principal's problem is denoted P'_n and is written as,

$$\text{Max } (q + (1 - q)m)[\alpha(\theta_1, e_1) - t_1] + (1 - q)(1 - m)[\alpha(\theta_2, e_2) - t_2]$$

s.t.

$$(2IR_1) \quad t_1 - \psi(e_1) \geqslant 0,$$

$$(2IR_2) \quad t_2 - \psi(e_2) \geqslant 0,$$

$$(2IC''_\gamma) \quad m = qz/[(1 - q)(P - z)].$$

It is clear that the constraints will be binding and production will be efficient.

Part (d). First we show that $\gamma > 0$ is optimal for P high enough. Given part (c), the principal's payoff under random audits differs from the first-best payoff only by m. However, ($2IC''_\gamma$) implies that $\lim_{P \to \infty} m = 0$. Therefore, it is higher than his payoff under second best for P high enough.

We now show that $\gamma = 1$ is not optimal for large values of P. Suppose $\gamma = 1$. Part (b) implies that $m = 1$. Also, ($2IR_2$) requires that $t_1 - \psi(\hat{e}_1) – P \geqslant 0$. Since the rent $t_1 - \psi(\hat{e}_1)$ is bounded due to our assumptions on $\alpha(\cdot)$ and $\psi(\cdot)$, constraint ($2IC_m$) is violated for large values of P. □

APPENDIX B

Proof of Proposition 1

$\gamma > 0$ and ($3IC_\gamma$) together imply $m > 0$, and $\gamma < 1$ and Proposition 0 (b) together imply $m < 1$. The constraints ($3IC'_\gamma$) and ($3IC'_m$) are the indifference conditions associated with a mixed strategy equilibrium, and using $\phi = (1-q)m/[q+(1-q)m]$, they are written as:

$$(3\mathrm{IC}'_m) \quad t_2 - \psi(e_2) = t_1 - \psi(\hat{e}_1) - \gamma P,$$

$$(3\mathrm{IC}'_\gamma) \quad (P-w)(1-q)m/[q+(1-q)m] = z,$$

which define γ and m in terms of the efforts and transfers. Replacing γ in the objective function, and γ and m in the (3IR) constraints, and solving for m to obtain ($3\mathrm{IC}''_\gamma$), the principal's problem is rewritten as **[P'_{NC}]**:

$$\text{Max } (q+(1-q)m)[\alpha(\theta_1, e_1) - t_1] + (1-q)(1-m)[\alpha(\theta_2, e_2) - t_2]$$

s.t.

$$(3\mathrm{IR}_1) \quad t_1 - \psi(e_1) \geqslant 0,$$

$$(3\mathrm{IR}_2) \quad t_2 - \psi(e_2) \geqslant 0,$$

$$(3\mathrm{IC}''_\gamma) \quad m = qz/[(1-q)(P-w-z)],$$

$$(\mathrm{CIC}) \quad w - \lambda P \geqslant 0.$$

It is clear that the (3IR) constraints are binding; otherwise the transfers can be reduced without violating any constraint. Thus, there is no rent in either state. Replacing t_i with $\psi(e_i)$ in the objective function, it is clear that efforts must be efficient. Then, ($3\mathrm{IC}'_\gamma$) implies that γ is identical to that in Proposition 0 (c).

The constraint (CIC) must also be binding. Given optimal efforts and transfers, the principal's payoff decreases with m, and ($3\mathrm{IC}''_\gamma$) implies that m increases with w. Finally, the binding (CIC) and ($3\mathrm{IC}''_\gamma$) implies that m is higher when collusion may occur, which in turn implies that the principal's payoff must also be lower. □

Proof of Proposition 2

Consider penalty levels such that $0 < \gamma < 1$ is strictly optimal if $\lambda = 0$, i.e., Proposition 0 (d) applies. We will show that there exists $\lambda(P) \in (0, 1)$ such that $0 < \gamma < 1$ is optimal if $\lambda < \lambda(P)$, and $\gamma = 0$ is optimal if $\lambda > \lambda(P)$.

Under the assumption on P, $\gamma = 1$ is never optimal. We will proceed by steps to show that for small λ random audits are optimal, and for high λ the second-best is optimal.

i) Using $\phi(1-\lambda)P = z$ and the binding (CIC) defines a function $m(\lambda)$ with the following properties:
$m(\lambda_1) \equiv 1$ where $0 < \lambda_1 < 1$, $m(0) > 0$, $m(\lambda)$ is continuous and strictly increasing for $\lambda \in [0, \lambda_1]$.

ii) For optimal efforts and transfers, the principal's payoff in problem [P'_{NC}] is continuous and strictly decreasing in m, and it approaches $[\alpha(\theta_1, e_1^*) - \psi(e_1^*)]$ as m tends to 1.

iii) We use Proposition 0 to show that the second best contract will dominate the best random audit contract if random auditing implies m close to 1. The

> second-best payoff must be higher than this limiting payoff $[\alpha(\theta_1, e_1^*) - \psi(e_1^*)]$ since the limiting payoff is feasible — set $e_1 = e_1^*$, $e_2 = \hat{e}_1$, and $t_1 = t_2 = \psi(e_1^*)$ — but not implemented when the second-best contract is chosen.

The proof is completed by noting that random auditing is strictly optimal for $\lambda = 0$. □

APPENDIX C

Proof of Proposition 3

Step 1. *Random internal and external audits ($0 < \gamma < 1$ and $0 < \delta < 1$) imply random shirking and collusion ($0 < m < 1$ and $0 < \mu < 1$).*

The constraints (EA-IC$_\gamma$) and (EA-IC$_\delta$) are indifference conditions (EA-IC$'_\gamma$) and (EA-IC$'_\delta$):

(EA-IC$'_\gamma$) ϕ [illegible] $\mu)(P^a - w) + \mu\delta(P^a + P^i - z_e)] - (1 - \phi$ [illegible] $z_e = z_i,$

(EA-IC$'_\delta$) $\phi^\mu(P^a + P^i) = z_e.$

(EA-IC$'_\delta$) implies that $m > 0$ and $\mu > 0$. It must be that $\mu < 1$, otherwise (EA-IC$'_\gamma$) and (EA-IC$'_\delta$) are both satisfied only if $z_i = 0$. Therefore, (EA-IC$_\mu$) and (EA-IR$_B$) imply that

(EA-IC$'_\delta$) $\delta = (P^a - w)/(P^a + P^i).$

The case $m = 1$ and $0 < \gamma < 1$ is strictly dominated by the second-best contract ($\gamma = 0$). To see this, use (EA-IC$'_\gamma$) and $m = 1$ in the objective function, and then the principal's problem is to maximize $\{\alpha(\theta_1, e_1) - t_1\}$ subject to (EA-IR$_1$). Clearly, the second-best yields a higher payoff and $\gamma > 0$ cannot be optimal. Therefore $0 < m < 1$, and using (EA-IR$_B$) we have

(EA-IC$'_m$) $t_2 - \psi(e_2) = t_1 - \psi(\hat{e}_1) - \gamma P^a.$

Step 2. *The optimal contract under random internal and external audits has efficient production when there is no shirking, no rent and $w = 0$.*

Using the four indifference conditions (EA-IC$'_\gamma$), (EA-IC$'_\delta$), (EA-IC$'_m$), and (EA-IC$'_\mu$), the principal's problem can be simplified as in sections 2 and 3 to show that production is efficient when there is no shirking and there is no rent. The only variable left to be determined is w. For optimal efforts and transfers, the principal's objective function is

$$(q + (1 - q)m)[\alpha(\theta_1, e_1^*) - \psi(e_1^*)] + (1 - q)(1 - m)[\alpha(\theta_2, e_2^* - \psi(e_2^*)],$$

where m is obtained by solving (EA-IC$'_\mu$), (EA-IC$'_\gamma$), and (EA-IC$'_\delta$). It is clear that the principal's payoff decreases with m. We will show that m increases with w, and therefore $w = 0$.

First, note that (EA-IC$'_\delta$) implies that $\mathrm{d}\mu/\mathrm{d}m < 0$. Then, use (EA-IC$'_\mu$) in (EA-IC$'_\gamma$) to get

(EA − IC$''_\gamma$) ϕ [illegible] $+ P^i) - \mu z_e] - (1 - \phi$ [illegible] $z_i/\delta,$

from which we obtain

$$\frac{dm}{d\delta} = \frac{z/\delta^2}{-\left[\frac{d\phi^{\mu}}{dm}(P^a + P^i - \mu z_e) - \phi^{\mu} z_e \frac{d\mu}{dm}\right] - z_e \frac{d\phi^{\mu}}{dm}} < 0,$$

where the inequality follows from $\mathrm{d}\phi/\mathrm{d}m > 0$ and the previous equation. Finally, from (EA-IC$'_{\mu}$) we know that δ increases with w.

Step 3. *Random internal and external audits are optimal.*

(i) First we prove by contradiction that a contract with $\delta = 1$ is dominated by the random audit contract. If $\delta = 1$, then $\mu = 0$ by (EA-IC$_{\mu}$), but then (EA-IC$_{\delta}$) requires that $\delta = 0$ which is a contradiction.

(ii) Next we show that a contract with $\gamma = 1$ is dominated by the random audit contract. By (EA-IC$_m$), $\gamma = 1$ is not optimal if P^a is large enough since the optimal efforts and transfers are bounded due to our assumptions on $\alpha(\cdot)$ and $\psi(\cdot)$.

(iii) Next we show that a contract with $\gamma = 0$ and $0 < \delta < 1$ is dominated by the random audit contract if z_i is small enough since there is more shirking in the first case. If $\gamma = 0$ and $0 < \delta < 1$ is optimal, the optimal contract is given by Proposition 0 but with cost of audit z_e instead of z_i. The probability of shirking is then

$$m^e = qz_e/(1-q)(P^a - z_e).$$

When $0 < \gamma < 1$ and $0 < \delta < 1$ is optimal, the conditions (EA-IC$'_{\mu}$), (EA-IC$'_{\gamma}$), and (EA-IC$'_{\delta}$) imply

$$m = \frac{q\left(\frac{z_i}{\delta} + z_e\right)}{(1-q)\left(P^a + P^i - \mu z_e - \frac{z_i}{\delta}\right)},$$

which shows that $m < m^e$ if z_i is small enough. Since the principal's payoff only differs by m between the two cases, our result follows.

(iv) Now we show that a contract with $\delta = 0$ is dominated by the random audit contract. Since $\lambda = 1$, by corollary 1 there can be no audit if $\delta = 0$, i.e., then the second-best contract is offered.

(v) Finally we show that the second-best contract is dominated by the random audit contract. If either penalty tends to infinity, (EA-IC$'_{\delta}$) implies that μm tends to zero. Since $w = 0$, (EA-IC$'_{\mu}$) implies that δ approaches one as P^a tends to infinity. Then (EA-IC$''_{\gamma}$) implies that ϕ^{μ} tends to zero as P^a becomes very large, which means m tends to zero. Therefore, the principal's payoff approaches the first-best payoff. □

Proof of Proposition 4

When doing comparative statics with the penalties, the optimal efforts and transfers do not change. Therefore, the principal's payoff changes only due to changes in m. We will first show that the decrease in m is greater for an increase in P^a compared to an increase in P^i. Since the principal's payoff is continuous in m, the limit results below complete the proof.

Solving for μ from (EA-IC$'_\delta$), and δ from (EA-IC$'_\mu$), and replacing in (EA-IC$'_\gamma$) gives us m

$$m = \frac{q(P^i - z_e)z_i + P^a(z_e + z_i)}{(1-q)(P^a + P^i - z_e)(P^a - z_i)}$$

Differentiating m with respect to P^a and P^i, we obtain

$$\frac{dm}{dP^a} = \frac{q(2P^a(P^i - z_e)z_i + P^i(P^i - z_e) + (P^a)^2(z_e + z_i)}{-(1-q)(P^a + P^i - z_e)^2(P^a - z_i)^2} < 0,$$

$$\frac{dm}{dP^i} = \frac{qP^a z_e}{-(1-q)(P^a + P^i - z_e)^2(P^a - z_i)^2} < 0,$$

and the difference between the two is

$$\frac{dm}{dP^a} - \frac{dm}{dP^i} = \frac{q(P^a + P^i)z_i}{-(1-q)(P^a + P^i - z_e)^2(P^a - z_i)^2} < 0.$$

Moreover, using L'Hopital's rule, the limit of m as P^a goes to infinity is zero, while the limit of m as P^i goes to infinity is $qz_i/(1-q)(P^a - z_i)$, which is identical to the probability of shirking when collusion is not allowed (section 2).

REFERENCES

Baiman, S.; Evans, J. and Noel, J., 1987, 'Optimal Contracts with a Utility-Maximizing Auditor,' *in Journal of Accounting Research*, vol. 25, no. 2, pp. 217–244.

Baron, D. and Besanko, D., 1984, 'Regulation, Asymmetric Information and Auditing,' *Rand Journal of Economics*, vol.15, pp. 47–70.

Bester, H. and Strausz, R., 2000, 'Imperfect Commitment and the Revelation Principle: The Multi-agent Case,' *Economics Letters*, vol. 69, no. 2, pp. 165–171.

Bester, H. and Strausz, R., 2001, 'Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case,' in *Econometrica*, vol. 69, no. 4, pp. 1077–1098.

Chartier, J., 2002, 'Accounting Fraud Rising,' mimeo, available at http://money.-cnn.com/2002/01/11/companies/acct_scandals/.

Che, Y-K., 1995, 'Revolving Doors and Optimal Tolerance for Agency Collusion,' *RAND Journal of Economics*, vol. 26, 378–397.

Faure-Grimaud, A.; Laffont, J-J. and Martimort, D., 1999, 'The Endogenous Transaction Costs of Delegated Auditing,' in *European Economic Review*, vol. 43, no. 4–6, 1039–1048.

Faure-Grimaud, A.; Laffont, J-J. and Martimort, D., 2002, 'Risk Averse Supervisors and the Efficiency of Collusion,' in *Contributions to Theoretical Economics*, vol. 2, no. 1, Article 5.

Garoupa, N., 1997, 'The Theory of Optimal Law Enforcement,' *Journal of Economic Surveys*, 11, 267–295.

Graetz, M.; Reinganum, J. and Wilde, L., 1986, 'The Tax-Compliance Game: Towards and Interactive Theory of Law Enforcement,' *Journal of Law, Economics and Organization*, vol. 2, 1–32.

Grout, P., *et al*. 1994, 'Auditor Professional Judgement: Implications for Regulation and the Law,' *Economic Policy: A European Forum*, 9, no. 19, October, 1994, 307–344.

Itoh, H., 1991, 'Incentives to Help in Multi-Agent Situations,' *Econometrica*, vol. 59, 611–636.

Kessler, A., 2000, 'On Monitoring and Collusion in Hierarchies,' *Journal of Economic Theory*, vol. 91, no. 2, 280–291.

Khalil, F., 1997, 'Auditing without Commitment,' *RAND Journal of Economics*, vol. 28, no. 4, 629–640.

Khalil, F. and Parigi, B., 1998, 'Loan Size as a Commitment Device,' *International Economic Review*, vol. 39, no. 1, 135–150.

Kofman, F. and Lawarrée, J., 1993, 'Collusion in Hierarchical Agency,' *Econometrica*, vol. 61, no. 3, 629–656.

Kofman, F. and Lawarrée, J., 1996, 'On the Optimality of Allowing Collusion,' *Journal of Public Economics*, vol. 61, 383–407.

Laffont, J-J. and Martimort, D., 1999, 'Separation of Regulators Against Collusive Behavior,' in *RAND Journal of Economics*, vol. 30, no. 2, 232–262.

Laffont, J-J. and Rochet, J-C., 1997, 'Collusion in Organizations,' *Scandinavian Journal of Economics*, vol. 99, no. 4, 485–495.

Laffont, J-J. and Tirole, J., 1993, *A Theory of Incentive in Procurement and Regulation*, (MIT Press).

Lambert-Mogiljansky, A., 1994, 'Corruption in Procurement–The Economics of Regulatory Blackmail,' research paper no. 1994:4RE, University of Stockholm.

Martimort, D., 2000, 'The Life Cycle of Regulatory Agencies: Dynamic Capture and Transaction Costs,' *Review of Economic Studies*, vol. 66, 929–948.

Picard, P., 1996, 'Auditing Claims in the Insurance Market with Fraud: The Credibility Issue,' in *Journal of Public Economics*, vol. 63, 27–56.

New York Times, 2000, June 16. 'Asleep at the Books: A Fraud that Went On and On and On,' Section C, page 1.

Shavell, S., 2004, *Foundations of Economic Analysis of Law* (Harvard University Press).

Strausz, R., 1997a, 'Collusion and Renegotiation in a Principal-Agent Relationship,' *Scandinavian Journal of Economics*, vol. 99, no. 4, 497–518.

Strausz, R., 1997b, 'Delegation of Monitoring in a Principal-Agent Relationship,' *Review of Economic Studies*, vol. 64, no. 3, 337–357.

Tirole, J., 1986, 'Hierarchies and Bureaucracies: On the Role of Collusion in Organizations,' *Journal of Law, Economics and Organizations*, vol. 2, 181–214.

Tirole, J., 1993, 'Collusion and the Theory of Organizations,' in Jean-Jacques, Laffont, *Advances in Economic Theory: Proceedings of the Sixth World Congress of the Econometric Society* (Cambridge: Cambridge University Press).

U.S. Congress, 2002, Pub. L. No. 107-204, 116 Stat. 745 (codified in scattered sections of 11, 15, 18, 28, and 29 U.S.C.A.).