CESifo Economic Studies, Vol. 53, 3/2007, 363–388 doi:10.1093/cesifo/ifm016
Advance Access publication 2 November 2007

Optimal Regulation of Auditing

Marco Pagano and Giovanni Immordino*

Abstract

We study regulation of the auditing profession in a model where audit quality is unobservable and enforcing regulation is costly. The optimal audit standard falls short of the first-best audit quality, and it is increasing in the riskiness of firms and in the amount of funding they seek. The model can encompass collusion between clients and auditors, arising from the joint provision of auditing and consulting services: deflecting collusion requires less ambitious standards. Finally, banning the provision of consulting services by auditors eliminates collusion but may not be optimal in the presence of economies of scope. (JEL Classification: G28, K22, M42).

Keywords: auditing, regulation, enforcement, collusion.

1 Introduction

The recent corporate scandals involving major companies (Enron, Worldcom, Qwest, Sunbeam, Parmalat, etc.) have highlighted that the regulation of auditing and its enforcement are key determinants of the reliability of corporate information. For many companies involved in corporate scandals, auditors failed to report any misbehavior or substantive inaccuracy. These audit failures have damaged auditors' reputation as independent experts and monitors of accounting information.

As a result of this loss of confidence, there has been a shift from self-regulation and litigation-based enforcement of audit rules towards government regulation and public-driven enforcement.[1] In the US, the Sarbanes-Oxley Act (SOX) of 2002 established the Public Company Accounting Oversight Board (PCAOB), which, under the oversight of the Securities Exchange Commission (SEC), will register public accounting firms, and establish rules for auditing, quality control, ethics, independence and other standards. Moreover, it will inspect accounting firms, carry out disciplinary proceedings and impose penalties. The SOX has also greatly increased the financial resources devoted to the activity of the SEC.

* Marco Pagano is affiliated with Università di Napoli Federico II, CSEF and CEPR, Via Cintia, 80126 Napoli, Italy, e-mail: mrpagano@tin.it and Giovanni Immordino is affiliated with Università di Salerno and CSEF, 84084 Fisciano (SA), Italy, e-mail: giimmo@tin.it

[1] Auditing rules apply to the conduct of auditors: they prescribe how audits must be conducted. In contrast, accounting standards apply to firms: they concern the reporting principles and procedures that firms can use.

A similar shift is under way in other countries. The United Kingdom moved away from self-regulation by widening the scope and powers of the Financial Reporting Council, and specifically by creating a new subsidiary board (the Professional Oversight Board for Accountancy), entrusted with oversight of the auditing and accountancy profession. Also, the Italian parliament is drafting new legislation to extend the powers of the national securities commission (CONSOB) on regulation and oversight of auditors' activity.

Now that the role of the public regulation of auditing is widely recognized, the natural question arises of what is the optimal design of such regulation. The job of auditors is to certify the reliability of accounting information, but in turn the reliability of this certification can vary depending on the quality standards set by regulators.[2]

These standards may concern the auditing procedure, the organization of the auditing firm, or both. For example, regulation can affect the auditing procedure by mandating external confirmation of the audited company's credits, and by calibrating the evidence required for such confirmation according to the credit's magnitude. While relatively small credits may be checked by a telephone call to the debtor or a fax, original documents may be required for credits whose existence and terms can affect the solvency of the audited firm. The importance of rules on external confirmation was highlighted by the recent Parmalat scandal, where massive fraud went undetected because of insufficient evidence on a major credit of the company.[3] The Parmalat scandal underscored also the importance of another procedural rule, which applies to the auditing of conglomerates by multiple auditors: when an auditor certifies the accounts of the group's holding company and other auditors verify those of its subsidiaries, a problem of "moral hazard in teams" arises unless the responsibility for the certification of the consolidated balance sheet lies with a single auditor.[4] But regulation

[2] In practice, regulators may differ in the definition of audit quality standards, as illustrated by the controversy between the US-based Financial Accounting Standard Board (FASB) and the UK-based International Accounting Standard Board (IASB). However, recently the differences between FASB and IASB are being reconciled via a sustained effort by these bodies to achieve consistency and comparability between their standards.

[3] On 6 March 2003, the auditing firm Grant Thornton accepted a copy of a fax sent by Bank of America as valid evidence of a EUR 3,6 bn credit and EUR 336 m cash held by Bonlat (a subsidiary of Parmalat), altogether worth 36 percent of Parmalat's debt and accounting for almost all the liquidity of the conglomerate. On 18 December 2003, the fax was revealed to be false. If the auditor had checked directly the existence of the credit with Bank of America, this fraud would have been revealed.

[4] Deloitte Touche was the group auditor and Grant Thornton dealt with subsidiaries, including the Bonlat offshore unit in the Cayman Islands, that is, at the centre of the scandal. Deloitte failed to detect the frauds because it took at face value the reports produced by Grant Thornton about Bonlat. If regulation had made Deloitte Touche the

can also affect the organization of auditing firms, for instance by providing guidelines regarding the system of quality control for audits and by setting standards of competence, independence and honesty of their employees.[5]

In this article, we show that the optimal design of auditing regulation depends on three main ingredients. First, the cost of enforcement, which includes both the necessary public funding (salaries of bureaucrats and judges, paperwork, investigations, etc.) and the compliance costs borne by audit firms and their customers. Second, the accountants' incentives to collude with their clients. Third, the economies of scope that may be reaped through the joint provision of auditing and consulting services to the same firm.

We characterize optimal regulation of auditing under the assumption that the quality of the information certified by auditors is unobservable, so that in the absence of regulation the equilibrium level of audit quality would be inefficiently low. To avoid this loss of informational efficiency and the implied misallocation of investment, the regulator can impose a minimum quality standard on auditors, but this choice must take enforcement costs into account. As a result, the optimal standard will fall short of the first-best audit-quality level, and must be lower, the less efficient the enforcement technology.

This baseline model assumes that the moral hazard problem lies only in the activity of auditors, while managers always seek a truthful report, since they are assumed to maximize shareholders' value. But the recent corporate scandals suggest that also the behavior of managers may be plagued by moral hazard, insofar as they want to go ahead with investment irrespective of its profitability, in order to extract private benefits from the firm's expansion. In this case, they will attempt to bribe auditors into producing positive reports under all circumstances. To do so, managers can award profitable consulting contracts to auditors, conditioning on receiving a favorable audit. The potential for collusion increases the auditors' incentives to misreport, so that more resources are required to enforce any given audit standard. Thus, when the managers of

only auditor responsible for the entire group, it would have raised its incentive to check the subsidiaries' accounts directly. Italy, the US and South Africa are among the few countries where this rule is absent, although legislation currently being passed in Italy will eliminate this deficiency. (We thank Roberto Tizzano for bringing this point to our attention.)

[5] Such rules on how to structure and operate auditing firms are not only present in the legislation of several countries, but also detailed in the standards issued by the International Auditing and Assurance Standards Board: see IAASB (2004).

client firms seek to corrupt their auditors, the regulator must optimally choose a less ambitious standard.

In this case, an additional regulatory tool is to sever the link between consulting and auditing activity, by forbidding auditors to provide consulting services, as indeed prescribed by the SOX. If this is the only way in which client firms may "bribe" their auditors, this policy would appear as a superior option to tampering with auditing standards. Indeed, in our model it would allow the regulator to leave the standard at the second-best level. However, this option is not necessarily superior once one takes into account that the profits arising from the joint provision of auditing and consulting are themselves part of social welfare.

Indeed, these profits have two conflicting effects on social welfare: a harmful effect, insofar as they raise the potential for collusion between auditors and their clients; and a beneficial effect, insofar as they reflect efficiencies of scope. We find that the regulator will want to allow bundling when the implied profits do not exceed a threshold that depends on the marginal efficiency of enforcement.

The structure of the article is as follows. Section 2 reviews the related literature and places the article in perspective, by comparing regulation with alternative mechanisms that can improve the informativeness of audits. In Section 3, we present the model, derive the first-best audit quality, and characterize the second-best audit standard to be chosen if audit quality is privately unobservable. In Section 4, we analyze the optimal regulation when firms collude with auditors. Section 5 considers how the design of regulation is affected by economies of scope arising from the joint provision of auditing and consulting services. Section 6 concludes the article.

2 Related literature

Our model is related to the microeconomic analysis of the auditor-firm relationship proposed by Dye (1993), where auditors can contribute to the efficient allocation of investment but the quality of their audits is unobservable, leading to a moral hazard problem. In contrast to Dye's model, however, in our setting this problem is not left to litigation between investors and accountants, but entrusted to regulation and its enforcement by public officials.[6] The optimal regulatory response must take into

[6] The assumption that enforcement of auditing standards is entrusted to litigation between investors and auditors is common to other papers in the accounting literature, such as Ewert (1999), who shows that under a negligence liability system for auditors vague standards may outperform precise standards in terms of auditing effort, quality of the financial statements, costs of enforcement and total cost of auditing.

account its enforcement cost, as in Immordino and Pagano (2007a). The result is a normative analysis of the regulation of the auditing profession, which takes into account the possible conflicts of interest of auditing firms, as well as the agency problems between managers and shareholders within client firms.[7]

Therefore, our article is relevant to the ongoing debate about the appropriate regulatory response to the recent corporate scandals. Our analysis focuses on regulation and public enforcement as the only device to temper agency problems in auditing and collusion between auditors and client companies. Private mechanisms have been suggested in the literature as alternative remedies to these problems: auditors' self-regulation assisted by litigation-based enforcement, reputational mechanisms, certification by intermediaries, financial statements insurance (FSI), whistleblowing, etc. But none of these mechanisms is indisputably superior to public regulation.

Indeed, as mentioned in the Introduction, the recent corporate scandals highlighted the weaknesses of self-regulation and private litigation as exclusive discipline devices for auditors. The reliability of audit companies as self-regulating entities has been tarnished by the increasing conflict of interest between their auditing role and their consulting role, as the share of consulting fees kept increasing in their revenues. This point is reminiscent of the result in DeMarzo, Fishman and Hagerty (2005) that a self-regulatory organization accountable to its members tends to choose laxer enforcement than customers would.

Of course, even in a system of public regulation, enforcement can be entrusted to litigation rather than to regulatory intervention, as postulated in our model. The limitation of litigation-based enforcement is that the costs of suing auditors may deter dispersed investors from taking action against violations, due to collective action problems. In contrast, a well-endowed and highly motivated public prosecutor can be very effective against financial fraud, as witnessed by the activity of New York State Attorney General Eliot Spitzer since 1999.

[7] Our model is also related to the seminal contribution by Leland (1979) on the benefits of minimum quality standards in markets with asymmetric information. Moreover, several contributions in industrial organization have analyzed the competitive effects of minimum quality standards: see, for instance, Ronnen (1991). However, our analysis differs considerably from this literature, because in our setting the minimum quality standard must take into account not only the incentives of producers (in our case, auditors) but also the possible agency problems on the demand side of the market (between shareholders and managers). These agency problems create possible collusion between the auditor and the manager against their common principal, i.e. shareholders. This element is absent in the existing literature on quality standards.

Reputation is another decentralized mechanism that might enhance the reliability of auditors, especially considering the limited number of active auditing firms, the repeated nature of their interactions with client firms and investors, and especially the large stakes represented by the auditors' equity base. Therefore, in principle this mechanism could be effective. However, it appears not to have deterred negligent or fraudulent behavior so extensive as to wipe out established companies such as Arthur Andersen. Even though the reasons why reputation has been ineffective are still unclear, its limitations suggest that it needs to be complemented by regulatory intervention.

Alternative mechanisms that could in principle be used to improve the reliability of auditors are intermediaries that "certify" the quality of their information (Lizzeri 1999) or the creation of a "financial statement insurance" (FSI) scheme, by which companies purchase insurance that protects investors against losses due to misrepresentation in financial reports and the insurer itself appoints and pays auditors (Dontoh, Ronen and Sarath 2004). Certification intermediaries would tend to realign the incentive of auditors towards truth-telling by creating a second layer of monitoring, while the FSI scheme would eliminate the conflict of interest arising from managers hiring auditors. But neither of these mechanisms is completely collusion-proof: both a certification intermediary and an insurance company providing FSI might collude with the client firm, unless they are themselves under the purview of a public enforcement mechanism.[8] Thus, rather than dispensing with public intervention, these mechanisms simply shift the need for such intervention to a higher layer. The FSI scheme suffers also from another problem: auditors appointed by outsiders may be unable to access sensitive data placed under management's control, and this lack of cooperation by management may prevent them from performing effective audits.

The danger of collusion between auditors and their clients suggests that "whistle-blowing" mechanisms might be particularly suited in this case, being designed to break collusion. Such mechanisms could reward individuals (e.g. employees) who report fraud by auditors and/or their clients, by entitling them to a portion of the penalty paid by the latter. Such schemes may deter collusion, as shown by Spagnolo (2004) in the context of antitrust regulation. But in our case whistle-blowing is unlikely

[8] Collusion can hinder the effectiveness of a FSI scheme in two ways. First, the fact that auditors are chosen and paid by the insurance company does not rule out that client firms might try to bribe them. Second, the client company might bribe the insurance company to underreport FSI premiums in exchange for higher premiums on other insurance contracts, and thereby mischievously appear in the eyes of investors as companies with high-quality financial statements.

to be effective, since the "whistle-blower" would hardly be able to document the *tacit* exchange of favorable audits against consulting contracts.[9]

This discussion underscores the importance of public regulation and enforcement as "residual" mechanisms to discipline the auditing profession.

3 The model

This section explains the rationale for regulation of auditing in a setting where auditing has informational value in raising new finance, as in Dye (1993). As a benchmark case, we first analyze a setting where the auditors' activity is observable and contractible, and the economy achieves the first-best outcome. We then examine what happens if investors cannot observe the level of effort that auditors invest in their task. This moral hazard problem in auditing implies that auditors will choose the minimal level of quality. Under our assumptions, the social cost of this moral hazard is that investors will allocate their funds less efficiently.

3.1 Informational value of auditing

Consider an economy with risk neutral agents and a continuum of firms. The representative company is managed in the interest of the shareholders, so that the manager's objective is to maximize its current value.

To continue operating, the company needs a cash injection (investment) of size I. Such refinancing being absent, the company is liquidated at a value that for simplicity is normalized to zero. The firm also needs to raise cash to pay for any fees F required by its auditors. Assuming that the required rate of return on new capital is standardized to zero, shareholders provide the needed cash infusion in exchange for shares that are worth at least $I+F$.[10]

Eventually, the company may turn out to be a success (state $s=H$) or a failure (state $s=L$). State H occurs with unconditional probability p, and state L with probability $1-p$. If the company is successful, its final value $\tilde{V}$ is V_H; if not, it is $V_L < I < V_H$. Thus, in the bad state it is not worth refinancing the company. Since there is a continuum of firms, p is also the fraction of successful firms. The initial shareholders and the manager are supposed to have no private information about the

[9] We thank Giancarlo Spagnolo for pointing this out to us.

[10] A smaller stake would violate their participation constraint. It is indifferent whether this cash infusion is contributed by the initial shareholders or by new shareholders. It is also immaterial whether the necessary cash is raised by issuing shares or risky debt.

company's future value.[11] So, absent any additional information, its market price P is the unconditional expectation of its final value, $\overline{V} = pV_H + (1-p)V_L$. We assume that $\overline{V} > I$, so that it is worth refinancing the company even if no information is gathered via an audit report.

However, an audit may still be worthwhile as it allows investors to condition the refinancing decision on an early signal of the firm's prospects. If the company is audited before it raises additional equity, its market price will reflect also the information certified by the audit. Auditors have a costly technology to elicit a signal that aids in distinguishing high-value from low-value firms, and relate this signal by filing a report r on the value of the firm. In practice, auditors assess only the reliability of the historical and prospective information provided by the company's accountants, and deliver this "filtered" information to investors who use it to evaluate the company.[12] As in Dye (1993), we collapse these two phases (the validation of accounting information and the evaluation made by the market) in a reduced-form process, by viewing the auditor's report as an assessment of the value of the company.

An auditing firm can choose the precision of its signal that it observes. For simplicity, the signal is assumed to be perfectly accurate when the state is H, while it may be inaccurate if the state is L. The auditor truthfully relates the signal in his report r.[13] Formally, the conditional probabilities of the auditor's report being correct are:

$$\begin{aligned} &\Pr(r = L \mid s = L, q) = q, \\ &\Pr(r = H \mid s = H, q) = 1. \end{aligned} \tag{1}$$

We assume that the accuracy q of the auditor's signal is influenced by the "quality" of the audit, as determined by the procedures adopted

[11] In many real-world circumstances, however, managers have private information about the firm's true value. In this case, auditors also perform the role of monitoring the manager's information. We analyze this alternative setting in a companion paper (Immordino and Pagano 2007b), where however the effort of auditors is assumed to be contractible, in contrast to the present setting. As shown in that paper, the introduction of asymmetric information has an important consequence for the role of auditors: an audit helps to monitor management, besides having an informational value as in the present model. This monitoring role of auditors arises to the extent that agency problems within the company make the information conveyed by managers not credible for shareholders.

[12] Auditors also solicit the production of additional information from the company's managers and accountants, when they are dissatisfied with its reliability. If met with refusal, they can threaten a "disclaimer of opinion".

[13] Even though the auditor cannot misreport the observed signal, he can choose to rely on a totally uninformative one. In this case, the signal will always be favorable to continuation of investment, as will be shown below.

(e.g. external confirmation of accounting data) and on the auditor's internal organization (e.g. personnel selection criteria). But better audit quality comes at a cost. Formally, the auditor chooses the audit quality $q \in (0,1)$ at a cost $C(q)$, which is continuous, increasing and convex in q, with $C(0) = 0$, $\lim_{q\to 0} C'(q) = 0$ and $\lim_{q\to 1} C'(q) = \infty$.

Using Bayes' rule, the probability that the company will succeed conditional on a good report is:

$$\Pr(s = H \mid r = H) = \frac{\Pr(s = H \cap r = H)}{\Pr(r = H)} = \frac{p}{p + (1-p)(1-q)}. \tag{2a}$$

while the probability that it will succeed conditional on a bad report is zero:

$$\Pr(s = H \mid r = L) = \frac{\Pr(s = H \cap r = L)}{\Pr(r = L)} = 0. \tag{2b}$$

The initial price of the company P equals the expectation of the company's value conditional on the auditor's report, which using the probabilities in Equations (2a) and (2b) can be written as

$$E(\tilde{V} \mid r = H) = \frac{V_H p + V_L(1-p)(1-q)}{p + (1-p)(1-q)} \geq \overline{V} \tag{3a}$$

if the report is favorable, or

$$E(\tilde{V} \mid r = L) = V_L \tag{3b}$$

if it is unfavorable. If no audit is carried out, the probabilities of the states H and L will be the unconditional ones, p and $(1-p)$, and the market price of the company will be equal to $\overline{V}$.

The shareholders' surplus from continuation, before netting out the audit fee F (if an audit is performed), $E(\tilde{V} \mid r) - I$, takes different values in each of these cases. With no audit, the surplus from continuation is $\overline{V} - I$, which is positive by assumption. So *a fortiori* the surplus is positive when the audit report is favorable. If instead the report is unfavorable, the surplus is $V_L - I < 0$. Therefore, the firm is refinanced only if no audit report is filed or if the report is favorable. It is liquidated if the report is unfavorable, so that in this case the surplus, conditional on the optimal investment decision, is zero.

Whenever an audit is commissioned, its cost F must be deducted from the shareholders' surplus. Let us denote by Δ_r the shareholders' surplus net of the audit fee F and conditional on a report r and on the optimal investment decision. This net surplus is $\Delta_H = E(\tilde{V} \mid r = H) - I - F$ and $\Delta_L = -F$, depending on the report filed by the auditor. The informational value of an audit is the difference between the

expected value of Δ_r with an audit and the surplus from continuation without an audit, $\overline{V} - I$:

$$\Omega(q) = \Pr(r = H)\Delta_H + \Pr(r = L)\Delta_L - (\overline{V} - I) = q(1-p)(I - V_L) - F. \tag{4}$$

This expression is increasing in the quality of auditing q, decreasing in the firms' quality p (the worse the pool, the more valuable is information), and increasing in the losses that would arise from investing in bad firms. The term $I - V_L$ is a measure of the potential misallocation of investment that can be prevented by auditors' information.

3.2 The unregulated outcome

If the audit fee F just equals the auditors' cost $C(q)$, i.e. if auditors make zero profits, then Expression (5) becomes the net *social* surplus (on a per-firm basis):

$$W(q) = q(1-p)(I - V_L) - C(q), \tag{5}$$

Indeed, since auditors earn zero profits, the entire net social surplus accrues to the shareholders.

The first-best outcome is obtained by maximizing the net social surplus $W(q)$. Given our assumptions about the auditor's cost function, $W(q)$ is concave and has an internal maximum where the marginal value of audit quality equals its marginal cost. This identifies the first-best quality value $q^{fb} \in (0,1)$:

$$(1-p)(I - V_L) = C'(q^{fb}). \tag{6}$$

Since the cost function $C(q)$ is convex, q^{fb} is decreasing in the quality of the pool and increasing in the potential misallocation of investment, just as the informational value of auditing.

If the audit quality is contractible, the first-best outcome emerges as the competitive market equilibrium. Firms' managers choose their "demand for audit quality" by maximizing the informational value of auditing, $\Omega(q)$. Auditors choose their "supply of audit quality" by maximizing their profit per audit, $F(q) - C(q)$, and make zero profits. The market-clearing price of an audit will then be the fee corresponding to the first-best level of audit quality, $F(q^{fb})$.

It is easy to show that if quality is observable, the first-best allocation coincides with the Bertrand equilibrium of the model, that is, the Nash equilibrium of an extensive-form game where auditors choose the quality q of the audit and a fee function $F(q)$. The strategy of auditor j is a choice of quality and fee, which is the best response to the qualities and fees chosen by competing auditors. The situation in which all firms choose the

first-best quality and price is a Nash equilibrium, since no firm can profitably deviate.[14]

If instead the audit quality is privately unobservable, then for any positive audit quality expected by investors, auditors have an incentive to choose a lower level and save the corresponding cost. As a result, the equilibrium audit quality is zero, the market price will equal the unconditional expectation $\overline{V}$, and an unprofitable firm will be more likely to continue operating.[15] So, in this case, there is a rationale for public intervention. To this, we turn in the next subsection.

3.3 Regulated auditing

The government sets an auditing quality standard q^*. This implies that auditors must choose a quality level at least equal to q^*. If they deviate, they are liable to pay a penalty l. The quality chosen by auditors is observable and verifiable at a cost by a regulator, who chooses also the amount of resources e devoted to enforcement, i.e. to detection of violators. The penalty is monetary and cannot exceed an upper bound,[16] denoted by l^*. This bound can be thought as the entire wealth of the auditing company, which is taken as exogenous in the context of the relationship with a specific client.[17]

Figure 1 illustrates the sequence of moves. First, nature chooses the state s. Second, the regulator chooses the audit standard q^*, the penalty l and the enforcement e. Third, auditors choose the quality level of their audit q charging the fee $F(q)$ and produce the corresponding report r. Fourth, bureaucrats enforce the standard by inspection, detecting

---

[14] If instead the auditor was a monopolist, he would extract all the informational surplus from its client by charging a fee equal to the client's valuation of its services, for any given value of q. To maximize his monopoly profits, he would then choose the first-best level of quality. So under monopoly audit quality would be efficient and the social surplus would be the same as under Bertrand competition. Only the distribution of this surplus would differ.

[15] Here we are assuming that the auditor's fee is not conditional on the *ex-post* accuracy of the report. Admittedly, such a contract could elicit a positive level of effort from the auditor, but it would not generally elicit the first-best auditing quality if auditors have limited liability. Moreover, in the auditing profession "contingent contracts cannot be used as an incentive device because auditors' code of professional ethics (rule 302) prohibits them from accepting incentive contracts for audit-related work" (Dye 1993, p. 888, fn. 5). This ban arises from efficiency reasons: auditors are hired by managers, and incentive contracts would provide easy ways to bribe auditors.

[16] Auditors' limited liability implies that the first-best cannot be achieved even via a more sophisticated design of the penalty or of enforcement, for instance by making them contingent on the accuracy of the auditor's report. (The point is analogous to that made in fn. 15.)

[17] Auditors' wealth can derive from interaction with previous customers and from the sale of non-audit services. In reality it may be impossible to confiscate the entire wealth of the auditor, due to the danger of "subversion of justice" (Glaeser and Shleifer 2003).

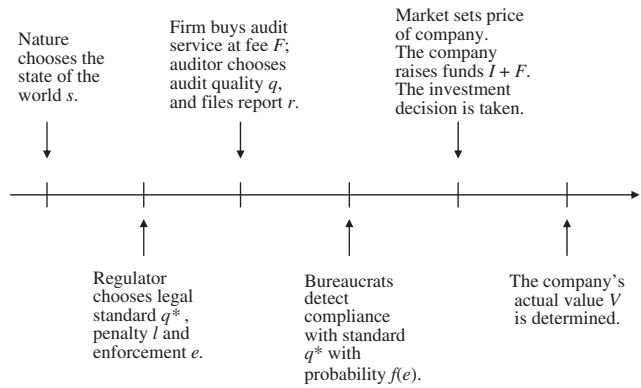

Figure 1 Time line

non-compliance with probability $f(e)$.[18] Next, the stock market sets the price P of the company at a level reflecting the actual report filed by the auditor and the perceived quality of auditing; depending on the report, shareholders may contribute equity to finance the company. Finally, the company's actual value is determined.

To enforce the audit standard, the regulator must back it up by a sufficiently high expected penalty L in case of noncompliance. The auditor's profit is:

$$\Pi = F(q) - C(q) - L, \tag{7}$$

where, the expected penalty L is the product of the probability $f(e)$ of detecting a non-complying auditor and the statutory penalty l. So the expected penalty is:

$$L = \begin{cases} f(e)l & \text{if } q < q^*, \\ 0 & \text{otherwise.} \end{cases}$$

The probability of detection is increasing and concave in the regulator's effort: $f'(e) > 0$, $f''(e) < 0$, that is, the enforcement technology has

[18] In our setting, the regulator commits to the probability of detection $f(e)$, by allotting resources e to enforcement activity. One can think of e as the salaries paid to the officials in the authority that oversees the application of audit standards: once hired, these detect violations with a probability given by their enforcement technology. We do not take explicitly into account the incentive problems that may arise within the enforcement agency. To the extent that such problems exist, they create agency costs that reduce the effectiveness of enforcement. In our setting, therefore, the introduction of such incentive problems in enforcement would be equivalent to a reduction of the probability of detection $f(e)$ for any given level of e, and would result in a reduction of the optimal audit standard.

decreasing marginal productivity. To ensure an interior solution for e and a detection probability comprised in the interval (0,1), we also assume $f(0) = 0$, $\lim_{e\to 0} f'(e) = \infty$, $\lim_{e\to\infty} f'(e) = 0$ and $\lim_{e\to\infty} f(e) \leq 1$.

Being benevolent, the regulator chooses the auditing standard q^*, the enforcement level e and the penalty l so as to maximize the social surplus from auditing quality *net* of the associated enforcement cost e, subject to the incentive-compatibility constraint of auditors. Formally, the problem is to maximize the second-best welfare level W^{sb}:

$$\max_{l,q^*,e} W^{sb}(q^*) = q^*(1-p)(I - V_L) - C(q^*) - e \tag{8}$$

subject to the incentive compatibility constraint:

$$F(q^*) - C(q^*) \geq F(q^*) - C(q) - f(e)l \text{ for any } q \neq q^*, \tag{9}$$

where, the auditor's fee F on both sides of the inequality corresponds to the prescribed audit quality expected by investors, while the cost C depends on the quality level actually chosen by the auditor.

As in Becker (1968), for any positive enforcement level it is optimal to set the penalty at the maximum feasible level:[19] $l = l^*$. To obtain the optimal enforcement level, we use the incentive compatibility constraint Equation (9) with equality, since the optimal policy requires this constraint to be binding. If not, the regulator could increase welfare by lowering enforcement e, for any given l^*. Next, notice that, in case of non-compliance, the auditor would optimally deviate to a zero quality level, since this would minimize his cost. Finally, since the detection probability $f(e)$ is monotonically increasing, it can be inverted to yield the optimal enforcement:

$$e(q^*) = f^{-1}\left(\frac{C(q^*)}{l^*}\right). \tag{10}$$

From the properties of the enforcement and audit technologies, it is immediate that the optimal enforcement e^* is an increasing and convex function of the audit standard q^*, and a decreasing function of the maximum penalty l^*.[20] The positive relationship between enforcement and audit standards highlights their complementarity: a more demanding audit standard invites non-compliance by auditors, so that it must be assisted by

[19] To see why, notice that if the penalty were set at a lower level, increasing it would enable the regulator to decrease enforcement e, while keeping L constant. The social surplus in the objective function would be unchanged but the enforcement cost would be lower, so that welfare would be higher.

[20] The first derivative of enforcement with respect to the standard is $e' = f'^{-1}(C(q^*)/l^*) \cdot C'(q^*)/l^* > 0$. Its second derivative is $e'' = f'^{-1}(C(q^*)/l^*) \cdot C''(q^*)/l^* + f''^{-1}(C(q^*)/l^*) \cdot (C'(q^*)/l^*)^2 > 0$.

more intensive monitoring by the regulator. Note that the stepwise penalty schedule just considered (by which all deviations below a target q^* are punished by the maximum penalty l^*) is not the only optimal penalty schedule.[21] Since these alternative penalty schedules are outcome-equivalent to the maximum penalty one, for simplicity we focus on the latter in the rest of the analysis.

Replacing the optimal enforcement Equation (10) into the objective function, the problem of maximizing Equation (8) can be rewritten as:

$$\max_{q^*} q^*(1-p)(I-V_L) - C(p^*) - e(q^*), \tag{11}$$

whose first-order condition is

$$(1-p)(I-V_L) = C'(q^*) + e'(q^*) \tag{12}$$

Under our hypotheses on the limiting behavior of the $C(q)$ and $f(e)$ functions, this optimality condition identifies an interior solution $q^* > 0$. More importantly, it implies that:

Proposition 1 (***Second-best audit standard***)

The optimal audit standard q^* is smaller than the first-best standard q^{fb}.

The proof of this proposition (and subsequent ones) is in the Appendix. The intuition for why the optimal standard is lower than the first-best level is simple: the regulator must take into account the resource cost of enforcing it. Note that, when regulation sets the audit standard at its second-best level q^*, the informational value of an audit is positive, since from Equation (5) it starts from zero for $q = 0$ and reaches its maximum at $q = q^{fb}$. Therefore, thanks to regulation, firms will want to hire an auditor.

It is interesting to explore how the optimal standard q^* varies depending on the parameters of the economy:

Proposition 2 (***Comparative statics***)

The optimal standard is decreasing in the fraction of successful companies p and increasing in the required investment I, in the efficiency of the auditing and in the efficiency of the enforcement technology.

---

[21] Indeed, an infinite number of distinct penalty schedules $l(q, q^*)$ can implement the standard q^* at the same enforcement cost $e(q^*)$ derived in Equation (10). These functions assign different penalties to different deviations by the auditor, in contrast to that considered above. All of these optimal schedules share the two following features. First, they all punish the deviation to the worst possible action by the auditor ($q = 0$) by the maximum penalty l^*: this is cost-efficient, as it minimizes enforcement effort. Second, these schedules are all weakly decreasing in the quality level q chosen by the auditor, so as to ensure that the incentive compatibility constraint Equation (9) is met: $l(q, q^*) \geq [C(q^*) - C(q)]/f(e(q^*))$. In other words, worse deviations are punished more harshly.

Intuitively, the audit standard should be more demanding if audits allow investors to pick the few winners in a bad pool, and/or if the audit cost is spread over a large investment. These are situations in which the social value of a reliable auditor is very high. Moreover, a country can afford higher standards if its auditors become more efficient in their job and/or regulators become better at monitoring them.

4 Auditors' conflict of interest and collusion with audited firms

As discussed in the Introduction section, one of the alleged sources of the recent corporate scandals has been the ability of managers to "buy" the acquiescence of auditing firms by exploiting the conflict of interest between their consulting arm and their auditing arm. Auditing firms can provide advisory services in the area of tax, accounting or management information systems as well as strategic advice, which for brevity shall be referred to as "advisory services". If managers behave opportunistically rather than in the best interest of investors, they can use the fees for the purchase of consulting services as "bribes" to buy the acquiescence of auditors to their objectives. To capture this point, we shall introduce the following new assumptions.

Instead of maximizing the firm's profits as assumed in previous sections, let us consider the situation where the manager maximizes the firm's investment, so as to reap a private benefit B stemming from the continued operation of the firm. This is a sum of money that he can appropriate at the shareholders' expenses without being punished. As a result, the shareholders' payoff from the investment in case of continuation is $\tilde{V} - I - B$.

Under these assumptions, managers will try to induce auditors to issue positive reports so as to allow the firm's continuation as often as possible. We assume that, owing to collective action problems, shareholders cannot dictate the appointment of auditors (otherwise, they would simply forbid the purchase of advisory services from the firm's auditor). Moreover, in contrast to the firm's manager, they do not know the cost of advisory service providers, but only the going price R of such services.

This asymmetry implies that shareholders cannot prevent opportunistic behavior by managers. However, we rule out direct bribes by the manager to the auditor via monetary transfers, on the ground that such openly illegal bribes would be detectable via "whistle-blowing", and therefore punishable by law enforcers. Direct monetary bribes are not allowed even when disguised as auditors' fees in excess of the competitive level: we

assume that this would be detectable by inspection of the company's accounts.[22]

Companies can purchase two services from auditors: auditing services at the fee $F(q) = C(q)$, as before; and advisory services at the competitive price R, which equals the cost of the marginal consulting company. The incumbent auditor has a cost advantage over the marginal provider of advisory services: he produces such services at zero cost, due to his acquaintance with the company and to economies of scope in the joint provision of auditing and advising.[23] So, if the incumbent auditor gets the consulting contract at the going price R for the provision of advisory services to the firm, he earns a rent R.

In the hands of an opportunistic manager, this rent could be used to buy the acquiescence of auditors. That is, the manager may offer a consulting contract at price R to the incumbent auditor conditional on obtaining a positive report $r = H$, irrespective of the true state of nature. Should the auditor refuse this offer, the manager will bargain with the auditor a discount for advisory services so as to extract at least part of the rent R. For simplicity, we assume that in this scenario the manager makes a take-it-or-leave-it offer and obtains the maximal discount R for the firm.[24]

This transaction obviously plays the same role of a bribe.[25] If he was to accept this "bribe", the auditor would optimally invest no resources in auditing, that is, he would choose zero quality since in our setting this will yield a positive report with certainty: $\Pr(r = H \mid q = 0) = 1$. In this

---

[22] If monetary bribes were allowed, the possibility of corruption would trivially arise even in the absence of the joint provision of auditing and consulting services. However, practitioners regard the joint provision of these services as the main source of auditors' involvement in corporate scandals (Crockett et al. 2003).

[23] The zero-cost assumption is made purely to save notation and implies no loss of generality: all that is needed is a potential cost advantage of the incumbent auditor over the competition.

[24] Strictly speaking, the manager—being completely opportunistic—is indifferent between hiring another consultant at price R or negotiating the price discount on behalf of the firm. We break this tie by assuming that he prefers to earn the price discount for the firm, for instance because he has a tiny equity stake and therefore draws a small financial benefit from the discount. For simplicity, we also assume that, in negotiating the price discount, the manager has all the bargaining power, since he makes a take-it-or-leave-it offer and therefore, extracts the entire surplus R on behalf of the firm. But our results generalize to less extreme bargaining solutions, where the manager is able to obtain consulting services from the incumbent auditor at the "price" αR, where $\alpha < 1$, rather than at zero prices. In this case, one must replace R with αR in the incentive constraint (14). Unless $\alpha = 0$, the results of this section are qualitatively unchanged.

[25] Admittedly, the potential for collusion may arise even within the market for auditing services, if this market is not perfectly competitive or if the incumbent has a cost advantage relative to the competition in the provision of auditing services. Also, in this case, auditors could be "bribed" by the threat of losing their rents. But in reality the rents appear to arise mainly from the joint provision of auditing and consulting services.

event, the incumbent auditor would earn $C(q) + R$, i.e., the sum of the cost saving in auditing $C(q)$ obtained by choosing zero quality instead of the statutory quality q and the cost saving R in the production of advising.

The manager's payoff from bribing the auditor derives from the implied increase in the probability of grabbing the private benefit: with the bribe, the firm continues with certainty, whereas without the bribe it would continue only if the auditor's report were positive, that is, with probability $\Pr(r = H)$. Therefore, the additional expected gain to the manager is $[1 - \Pr(r = H)]B$.[26]

This changes in two ways the regulator's problem of choosing the optimal audit standard and the associated enforcement spending. On one hand, the cost efficiency of the incumbent auditor increases social welfare. On the other hand, the implied rent can be used by the firm's manager to bribe the auditor. Formally, the regulator's objective function must be modified by including the cost savings R of the incumbent auditor:

$$\max_{l,q^c,e} \hat{W}(q^c) = q^c(1-p)(I - V_L) + R - C(q^c) - e, \tag{13}$$

where q^c denotes the standard chosen by the regulator when there is the danger of collusion. The new symbol $\hat{W}$ indicates that the social welfare function has been modified to include also the efficiency gain R. Instead, the presence of private benefits B does not affect the expression for social welfare, since they are a pure transfer.

The incentive compatibility constraint of the auditor must be rewritten to take into account that managers can use the rent R to corrupt incumbent auditors:

$$F(q^c) - C(q^c) \geq F(q^c) + R - f(e)l^*. \tag{14}$$

Now regulation and enforcement must be designed also to deter the auditor from accepting the bribe. Going through the same steps as in the previous section, one finds that now the optimal enforcement level is:

$$e_c(q^c) = f^{-1}\left(\frac{(C(q^c) + R)}{l^*}\right). \tag{15}$$

This expression, which is increasing and convex in q^c, shows that the enforcement necessary to uphold a given audit standard q^c exceeds the second-best level identified in Equation (10), because the possibility of obtaining the rent R from consulting raises the enforcement activity

[26] Therefore, the total surplus available to the manager and the auditor is the sum of the auditor's cost savings $C(q) + R$ and the manager's additional private benefits $\Pr(r = L)B$. Since monetary transfers are not allowed, they cannot split their joint surplus in any way that differs from the one considered in the text.

needed to prevent collusion. The optimal audit standard is correspondingly lower: intuitively, the potential bribe raises the cost of enforcement, inducing the regulator to choose a less ambitious standard. The larger is the rent R from advisory services, the greater the potential for corruption, and therefore the less ambitious the auditing standard must be:

Proposition 3 (Audit standard with conflict of interest)

If the rent R from advising is positive, then the optimal standard $q^c(R)$ is lower than the second best q^* for $R > 0$ and it is decreasing in R.

Note that the conflict of interest may induce the regulator not to intervene: the rent R increases discretely the incentive to deviate and thereby in certain cases the maximal value of welfare attainable with intervention would fall below zero: $\hat{W} < 0$ for any $q^c > 0$. In this case, the regulator will rather choose $q = 0$, in contrast with the second best.

One might suggest that an even better policy would be to eliminate the conflict of interest at its root, by severing the link between auditing and advisory activity. In this case, the optimal accounting standard would increase to the second-best level q^*. However, this does not necessarily increase welfare: even though auditing standards would increase, society would forgo the cost savings arising from the joint provision of auditing and advisory services. The next section analyzes under which circumstances a ban on such joint provision is justified.

5 Conflict of interest versus efficiency gains from non-audit services

Now, we consider the case in which the regulator, besides setting an auditing standard and an enforcement level, may forbid the provision of advisory services by auditors. Indeed, this is one of the provisions contained in the SOX. SOX prohibits all registered public accounting firms from providing audit clients, contemporaneously with the audit, certain non-audit services including internal audit outsourcing, financial-information-system design and implementation services and expert services. These scope-of-service restrictions go beyond previous SEC independence regulations. In addition, all other services, including tax services, are permissible only if pre-approved by the issuer's audit committee and all such pre-approvals must be disclosed in the issuer's periodic reports to the SEC. Similar reforms have recently been approved also in other countries. For instance, Italy's 2005 "Law for the Protection of Saving and Financial Market Regulation" (Article 18) contains a prohibition that is very similar to that contained in SOX.

If regulation allows bundling of auditing and advisory services, society reaps the efficiency gain arising from economies of scope, but opens the door to the managers' attempt to "bribe" their auditor. If instead the regulator forbids bundling of these services, it forgoes the efficiency gain R arising from economies of scope. If this strategy is chosen, the problem reverts to that of Section 3, and the resulting auditing standard becomes simply the second-best level q^*, derived at the end of Section 3.

Which of these strategies maximizes social welfare? The answer depends on two factors: (i) the magnitude of the rents R from consulting and (ii) the marginal efficiency of enforcement. Indeed, an increase in R has two conflicting effects: an increase in the potential for collusion between manager and auditor and an efficiency gain from auditors' economies of scope.

If the marginal efficiency of enforcement is sufficiently high ($e'(q^*) > C'(q^*)$), the negative effect of an increase in R is dampened by the law, so that the positive effect from the efficiency gain dominates. But this holds only for relatively low values of R: once the rents from consulting that can be used to bribe consultants become too high, the negative effect may dominate. This is illustrated in the top panel of Figure 2. The solid curve represents the welfare level (drawn for a concave welfare function) when the regulator allows the joint provision of auditing and consulting services, $\hat{W}(q^c; R)$, whereas the flat line indicates the welfare level associated with the prohibition of bundling, that is, the second-best welfare level $W^{sb}(q^*)$. Below the threshold $\tilde{R}$, the solid curve lies above the flat line, so that it is optimal to allow the joint provision of auditing and consulting services. Above this threshold, instead, it is optimal to forbid bundling. Note that the threshold may be unbounded, in which case it is never optimal to forbid bundling.

The bottom panel of Figure 2 shows how the optimal audit standard changes in response to different values of the rent R. The audit standard must be set at the second-best level q^* when $R = 0$, decreases as R increases (as shown by Proposition 3) up to the threshold $\tilde{R}$, and reverts to the second-best level q^* for even larger R. Intuitively, in the region where bundling is allowed, the standards must be gradually mellowed, as R increases in order to accommodate the increasing severity of the conflict of interest. This process stops as soon as the economy enters the region where it is optimal to forbid bundling, thereby removing the conflict of interest at its root.

But also another case is possible, which is illustrated by the dashed line in Figure 2: if the marginal efficiency of enforcement is sufficiently low ($e'(q^*) < C'(q^*)$), then the welfare level $\hat{W}(q^c; R)$ that obtains allowing the joint provision of auditing and consulting (the dashed curve) lies entirely below the second-best welfare level $W^{sb}(q^*)$ (the flat line). In this case,

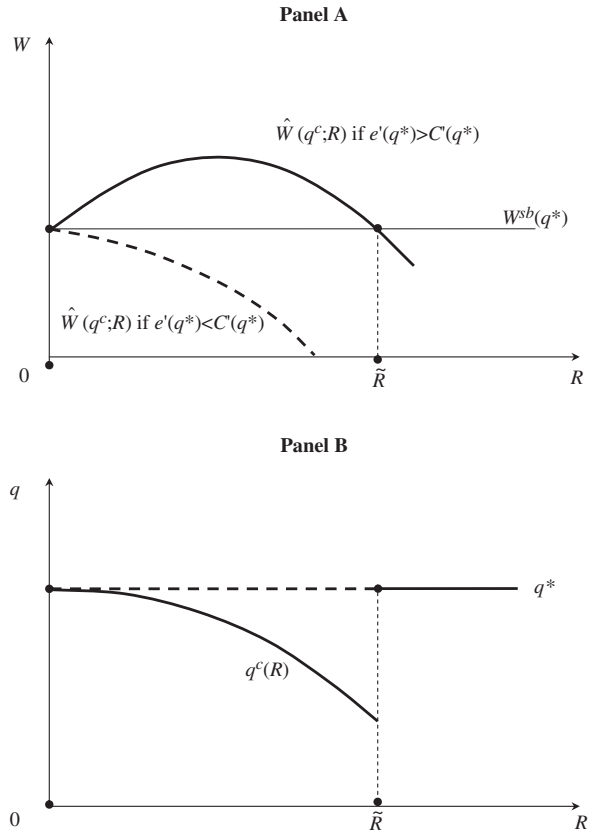

Figure 2 Social welfare W (panel A) and optimal auditing standards q (panel B) as functions of the rents from consulting R

bundling should always be forbidden, and the audit standard should always be set at its second-best level q^* for any R. In this case, the legal system is so inefficient that even a very small bribe to the auditor is too costly to deter.[27]

These points are summarized in the following proposition:

Proposition 4 (Interdiction of bundling and audit standards)

(1) If $e'(q^*) > C'(q^*)$, then there exists a possibly unbounded $\tilde{R} > 0$ such that for $R \leq \tilde{R}$, it is optimal to allow bundling of auditing and

[27] In the proof of Proposition 4 we also show that in the borderline case where $e'(q^*) = C'(q^*)$, social welfare is the same in the two regimes, so that the regulator is indifferent between allowing and forbidding bundling for any R.

consulting, and to forbid it for $R > \tilde{R}$. The optimal audit standard is $q^c(R)$ for $R \leq \tilde{R}$ and q^* for $R > \tilde{R}$.

(2) If $e'(q^*) < C'(q^*)$, then it is optimal to forbid bundling of auditing and consulting and set the audit standard equal to q^* for any R.

One may wonder whether, rather than banning the joint provision of audit and non-audit services, the regulator may not want to limit its intervention simply to mandating disclosure of information about the identities of the auditor and the consultant of the firm, trusting that shareholders will intervene to forbid the joint provision of the two types of service in the regions where this is inefficient. However, in our setting this type of mandatory disclosure would not prevent collusion between managers and auditors, due to the assumption of collective action problems among shareholders. Indeed, starting from 2000, the SEC had already mandated firms to disclose audit and non-audit fees paid to their auditors, hoping to provide investors with transparent information about the auditor-client relationship. The approval of SOX in 2002 indicates that this disclosure was not deemed sufficient to achieve the desired degree of auditors' independence.

As explained above, the results of Proposition 4 rest on the balance between the two conflicting effects of the rent R earned in the provision of non-audit services: on the one hand, the easier capture of auditors by managers; on the other, the efficiency gains from economies of scope. Various papers have attempted to measure empirically, which of these two effects dominates. Some event studies show that unexpected increases in non-audit fees have a negative or insignificant effect on the abnormal returns of companies (Chaney and Philipich 2002; Frankel, Johnson and Nelson 2002; Ashbaugh, Lafond and Mayhew 2003). Other studies have instead uncovered evidence of economies of scope. Some of them have done so by showing that audit fees are positively correlated with non audit-fees (Simunic 1984; Palmrose 1986; Parkash and Venable 1993). The study by Antle et al. (2006) shows that this positive correlation survives even after controlling for the joint determination of audit and non-audit fees as endogenous variables. In a recent paper, Cho, Han and Brown (2006) use a different empirical strategy: they investigate how stock market prices respond to accounting information and test whether the responsiveness of prices change depending on the ratio of audit to non-audit fees paid the company. They find that the price relevance of accounting information is greater in companies where non-audit fees are comparably larger, which can be seen as further evidence in favor of economies of scope in the activity of auditors.

6 Conclusion

The recent corporate scandals have highlighted the need for tighter regulation of the audit profession. However, once it is recognized that the enforcement of such regulation is costly, three important lessons can be drawn concerning the optimal standards to be imposed on auditors.

First, audit quality standards must be based on a cost-benefit analysis of audit activity. On the cost side, they must be less ambitious in economies that have less efficient enforcement. On the benefit side, they must be tighter in economies where the fraction of bad investments is larger or companies seek more external funding.

Second, regulatory standards must be less ambitious when auditors can collude with the managers of client companies at the expense of shareholders, because deflecting the potential for collusion requires more intensive—and therefore costlier—enforcement.

Third, if client firms may "bribe" their auditors by offering them generous consulting contracts, regulators can eliminate the source of collusion by forbidding auditors to provide consulting services. This policy prescription is not always warranted since the joint provision of auditing and consulting services may also generate socially valuable economies of scope. Banning such joint provision is socially inefficient if the associated conflict of interest is not too acute, and if enforcement is so efficient as to keep this conflict in check at relatively low cost for society.

References

Antle, R., E.A. Gordon, G. Narayanmoorthy, et al. (2006), "The Joint Determination of Audit Fees, Non-audit Fees, and Abnormal Accruals", *Review of Quantitative Finance and Accounting* **27(3)**, 235–66.

Ashbaugh, H., R. LaFond and B.W. Mayhew (2003), "Do Nonaudit Services Compromise Auditor Independence? Further Evidence", *The Accounting Review* **78**, 611–39.

Becker, G.S. (1968), "Crime and Punishment: An Economic Approach", *Journal of Political Economy* **76**, 169–217.

Chaney, P.K. and K.L. Philipich (2002), "Shredded Reputation: The Cost of Audit Failure", *Journal of Accounting Research* **40**, 1221–45.

Cho, S.-Y., J. Han and K.Fr Brown (2006), "Do Nonaudit Services Enhance Value? Evidence from the Capital Markets", SSRN Working Paper, available at http://ssrn.com/abstract=930556.

Crockett, A., T. Harris, F. Mishkin, et al. (2003), "Conflicts of Interest in the Financial Service Industry: What Should We Do About Them?".

Geneva Reports on the World Economy, 9. Centre for Economic Policy Research, available at http://www.cepr.org/pubs/books/cepr/booklist.asp?cvno=P163.

DeMarzo, P.M., M. Fishman and K.M. Hagerty (2005), "Contracting and Enforcement with a Self-regulatory Organization", *Review of Economic Studies* **72**, 687–706.

Dontoh, A., J. Ronen and B. Sarath (2004), "Financial Statements Insurance, New York University". SSRN Working Paper, available at http://ssrn.com/abstract=303784.

Dye, R.A. (1993), "Auditing Standards, Legal Liability, and Auditor Wealth", *Journal of Political Economy* **101(5)**, 887–914.

Ewert, R. (1999), "Auditor Liability and the Precision of Auditing Standards", *Journal of Institutional and Theoretical Economics* **155**, 181–206.

Frankel, R.M., M.F. Johnson and K.K. Nelson (2002), "The Relation Between Auditor's Fees for Nonaudit Services and Earnings Management", *The Accounting Review* **77**, 71–105.

Glaeser, E.L. and A. Shleifer (2003), "The Rise of the Regulatory State", *Journal of Economic Literature* **41(2)**, 401–25.

IAASB (2004), "Quality Control for Firms that Perform Audits and Reviews of Historical Financial Information, and Other Assurance and Related Services Engagements", *International Standard on Quality Control* **1**.

Immordino, G. and M. Pagano (2007a), "Legal Standards, Enforcement and Corruption", CSEF Working Paper No. 98. <http://www.dise.unisa.it/WP/wp98.pdf>, last accessed October 2007.

Immordino, G. and M. Pagano (2007b), "Corporate Fraud, Governance and Auditing", unpublished manuscript, June.

Leland, H. (1979), "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards", *Journal of Political Economy* **87(6)**, 1328–1346.

Lizzeri, A. (1999), "Information Revelation and Certification Intermediaries", *RAND Journal of Economics* **30(2)**, 214–31.

Palmrose, Z.-V. (1986), "The Effect of Non-audit Services on the Pricing of Audit Services: Further Evidence", *Journal of Accounting Research* **24(2)**, 405–11.

Parkash, M. and C.F. Venable (1993), "Auditee Incentives for Auditor Independence: The Case of Non-audit Services", *The Accounting Review* **68**, 113–33.

Ronnen, U. (1991), ''Minimum Quality Standards, Fixed Costs, and Competition'', *Rand Journal of Economics* **22(4)**, 490–504.

Simunic, D.A. (1984), ''Auditing, Consulting and Auditor Independence'', *Journal of Accounting Research* **22(Autumn)**, 679–702.

Spagnolo, G. (2004), *Divide et Impera: Optimal Leniency Programmes*, CEPR Discussion Paper No. 4840, December.

Appendix

Proof of Proposition 1. To show that $q^* < q^{fb}$, compare condition (12) to the first-best condition (6): the right-hand side of (12) is larger, implying the result by the convexity of $C(q)$ and $e(q)$. ■

Proof of Proposition 2. To help the reader's intuition, this proof is provided graphically rather than algebraically. Figure A1 illustrates the optimal audit standard chosen by a regulator. The convex function $e(q^*)$ shows the minimum enforcement required for each audit standard, from Equation (10). The government's preferences are described by a field of concave, upward-sloping social indifference curves, from the properties of the welfare function W: their slope $(1-p)(I-V_L)-C'(q)$ is positive for quality levels lower than the first-best, and is decreasing by the convexity of $C(q)$.

The optimal values of e and q^* are at the tangency between the lowest indifference curve and the $e(q^*)$ function. Consider an increase in the required investment I or a reduction in the fraction of good firms p: both of these increase the marginal value of the audit quality. This results in an increase in the optimal standard, because the social indifference curves become steeper. Similarly, higher efficiency of enforcement decreases the slope of the $e(q^*)$ function and thus raises the optimal standard. By the

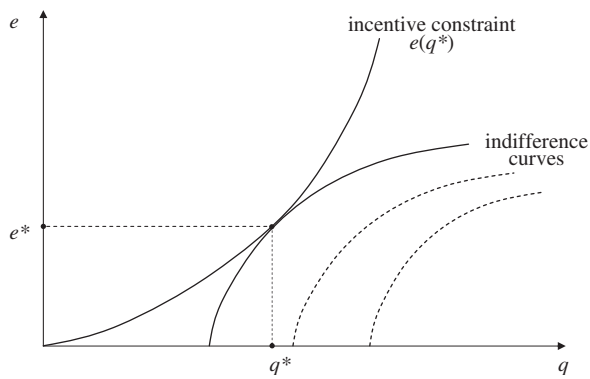

Figure A1 Equilibrium auditing standards q and enforcement level e

same token, a greater cost efficiency of auditors increases the optimal standard, since the social indifference curves become steeper and the $e(q^*)$ function flatter. ■

Proof of Proposition 3. The first-order condition for the regulator's problem now is:

$$(1-p)(I-V_L) = C'(q^c) + e_c'(q^c), \tag{A1}$$

The right-hand side of (A1) exceeds that of condition (12), while their left-hand sides are identical. By the convexity of $C(q)$ and $e(q)$, this implies $q^c < q^*$. By differentiating the first-order condition (A1) and Equation (15), we obtain:

$$\frac{\mathrm{d}q^c}{\mathrm{d}R} = -\frac{f''^{-1}((C(q^c)+R)/l^*)C'(q^c)/l^{*2}}{C''(q^c)+e_c''(q^c)} < 0.$$

This shows that q^c is decreasing in R. ■

Proof of Proposition 4. To maximize social welfare, the regulator has two alternative strategies: allowing the joint provision of auditing and consulting services or forbidding it. Let the welfare level that obtained in the first case be $\hat{W}(q^c; R)$. Instead, the welfare level associated with the prohibition of bundling is the second-best welfare level $W^{sb}(q^*)$.

We compare these two strategies, for alternative values of R. Note that for $R=0$, the welfare associated with the two strategies is the same by definition: $\hat{W}(q^c; R) = W^{sb}(q^*)$. While W^{sb} is independent of R, $\hat{W}(q^c; R)$ can either increase or decrease depending on the derivative:

$$\frac{\partial \hat{W}(q^c; R)}{\partial R} = 1 - f'^{-1}\left(\frac{C(q^c)+R}{l^*}\right)\frac{1}{l^*}. \tag{A2}$$

If this expression is negative for $R=0$, it remains negative also for any positive R. If it is positive, it becomes negative for R sufficiently large. To see this, note that using Equation (15), one can rewrite (A2) as:

$$\frac{\partial \hat{W}(q^c; R)}{\partial R} = \frac{1}{C'(q^c)}\left[C'(q^c) - e_c'(q^c)\right]. \tag{A3}$$

As R increases, this expression decreases: from Proposition 3, q^c decreases as R increases, and therefore $C'(q^c)$ decreases. From the first-order condition (A1), the sum $C'(q^c) + e_c'(q^c)$ must stay constant, implying that $e_c'(q^c)$ increases in R. As a result, the expression in square brackets in (A3) decreases in R, so that the derivative (A3) becomes negative for R sufficiently large.

This argument implies that if the derivative (A3) is positive at $R=0$, that is, $C'(q^*) - e'(q^*) > 0$, then $\hat{W}(q^c; R)$ exceeds W^{sb} for low R. So there exists

a (possibly unbounded) value $\tilde{R} > 0$ such that $\hat{W}(q^c; \tilde{R}) = W^{sb}$. If instead, the derivative (A3) is negative at $R = 0$, then $\hat{W}(q^c; R)$ is always smaller than W^{sb}. In this case, $\tilde{R} = 0$.

In the borderline case where $e'(q^*) - C'(q^*)$, the derivative is zero. Since social welfare is the same in the two regimes for $R = 0$, then also with bundling it must be the same and equal to W^{sb} for any R. ■