The Effects of Restricting Nonaudit Services on Audit Quality and Audit Value

Derek K. Chan
The University of Hong Kong

Nanqin Liu
University of Macau

ABSTRACT: This paper examines the economic consequences of prohibiting auditors from providing various nonaudit services (NAS) to their public company audit clients. In our model, auditors are strategic in choosing both effort and audit report, and the joint provision of audits and NAS to the same client creates synergies. We show that although restricting NAS can restore full auditor independence, it adversely influences the audit effort. Furthermore, we show that this regulatory change has unintended negative consequences on audit quality and audit value particularly when the auditor's legal liability is sufficiently small.

JEL Classifications: M42, M48

Keywords: *Sarbanes-Oxley Act; financial reporting; auditor effort; auditor independence; nonaudit services.*

Derek K. Chan gratefully acknowledges the Research Grant Council of the Hong Kong Government for financial support (HKU 74074/05H).

Derek K. Chan (corresponding author), The University of Hong Kong, Faculty of Business and Economics, Accounting and Law Area, Hong Kong, China. Email: dkchan@.hku.hk.

Nanqin Liu, University of Macau, Faculty of Business Administration, E22 Avenida da Universidade, Taipa, Macau, China. Email: nanqinliu@um.edu.mo.

I. INTRODUCTION

The restriction of the joint provision of nonaudit services (NAS) has been a controversial issue for worldwide policy makers, practitioners, and scholars. In the United States, Section 201 of the Sarbanes-Oxley Act (SOX) of 2002 prohibits auditors from providing many types of NAS to their public company audit clients.[1] The European Union (EU) introduced similar restrictions in 2016. Proponents of NAS restriction claim that the joint provision of audit and NAS to the same client may lure the auditor into compromising his/her opinion in the hope of grabbing the lucrative NAS businesses. This can impair auditor independence, reduce auditors' incentives to exert effort to gather audit evidence — as it will not be relied on anyway — and thus damage audit quality, which reflects the reliability of audit reports, and ultimately lead to bad investment decisions based on these reports.[2] By contrast, opponents argue that restricting NAS may damage synergies (or economies of scope) between audit and NAS, and thereby reduce audit effort, audit quality, and audit value.[3]

This paper sheds some light on the implications of this regulatory change. We study how a restriction of NAS would affect auditors' decision to exert effort to collect audit evidence (i.e., audit effort decision) and their willingness to faithfully express an opinion that reflects

---

[1] Section 201 specifically identifies nine types of NAS that cannot be provided by incumbent auditors, such as appraisal or valuation services, contribution-in-kind reports, actuarial services, management functions or human resources, broker or dealer, investment adviser, or investment banking services, legal services and expert services unrelated to the audit. There are still exemptions of NAS that auditors can jointly provide with audit service if audit committees of the issuers review and approve the NAS. A typical example is tax service.

[2] Surveys and experimental studies show that NAS performance affects financial statement users' perception of audit firm independence (Lowe et al. 1999; Swanger and Chewning 2001). By contrast, empirical findings on the association of joint provision of audit and NAS and impaired auditor independence are reported as inconclusive and conflicting (DeFond et al. 2002; Ashbaugh et al. 2003; Antle et al. 2006; Ruddock et al. 2006). Similarly, prior studies obtain mixed results regarding the relationship between NAS fees and output-based audit quality proxies (Frankel et al. 2002; Chung and Kallapur 2003; Raghunandan et al. 2003; Kinney et al. 2004; Robinson 2008).

[3] Simunic (1984) calls these synergies "knowledge spillovers", suggesting that the joint provision of audit and NAS may give rise to knowledge spillovers (from auditing to NAS and/or vice versa) and lead to economic rents. Arruñada (1999) points out that the joint provision of audit and NAS may raise the technical quality of auditing by broadening the auditor's knowledge of the client's business. Wu (2006) argues that auditors who also provide NAS to their clients may gain competitive advantages over their rivals. Prior studies (Simunic 1984; Davis et al. 1993; Chan et al. 2012) find evidence consistent with knowledge spillovers by showing a positive association between audit and nonaudit fees.

the evaluation of what has been discovered during the audit (i.e., auditor independence, or more specifically, the audit report decision). Through these effects, we can also examine the consequential impacts on audit quality (i.e., the joint probability of detecting a misstatement, if one exists, and reporting it) and the value of the audit (i.e., the sum of the assurance value and the synergy of the audit minus the audit cost).

We develop a simple model that examines the interactions among audit and NAS, legal damages for audit failures, and auditors' effort and report decisions. In our model, an auditor obtains client-specific knowledge from performing an audit that can create synergies if s/he jointly provides NAS. The auditor can obtain the NAS business and materialize the corresponding economic rents in the form of a consulting fee only if the auditor issues an unmodified report. The auditor is strategic in choosing both effort and report. Specifically, the auditor first exerts effort to collect audit evidence to identify whether the client is a potential recipient of a modified opinion (i.e., the auditor disagrees with the firm's original favorable report). The audit effort also increases the probability of obtaining the economic rents from the provision of NAS to the same client. The auditor then decides whether to faithfully report or suppress the unfavorable audit evidence (if any exists) by considering the following economic trade-offs. On the one hand, the auditor will lose the opportunity to earn economic rents from the consulting business upon issuing a modified opinion. On the other hand, failing to issue a modified opinion exposes the auditor to lawsuits. When making the report decision, the auditor compares the expected liability loss from an audit failure with the loss of the future economic rents, and may or may not report faithfully based on the audit evidence and the client-specific knowledge.

Prior to the NAS regulation, client firms could hire their auditors for both audit and NAS. The NAS regulation restricts auditor-provided NAS. We show that limiting the scope of NAS can improve auditor independence ex post, and a sufficiently strict restriction of NAS restores full auditor independence. However, restricting NAS also decreases audit effort ex ante. Given that audit effort increases the probability of obtaining the economic rents through providing NAS to the same client, restricting NAS has a *direct* effect that reduces the auditor's incentive to work hard by decreasing the marginal benefit of audit effort.

We also study the effects of NAS restriction on audit quality and audit value. Restricting NAS sacrifices some synergies between audit and NAS, and thus directly reduces the value of the audit. Restricting NAS also generates opposite effects on audit quality, which in turn affect the assurance value of the audit: although restricting NAS restores full auditor independence ex post and thus ensures faithful audit reports, it also adversely influences the audit effort ex ante and thus decreases the probability that the auditor finds unfavorable audit evidence. We show that the effect of lowering audit effort overwhelms the effect of improving auditor independence if the auditor's legal liability is sufficiently small. The key is that the auditor weights the expected consulting profit against the expected liability loss to determine whether to suppress the unfavorable audit evidence. Taking the auditor's legal liability as exogenous, a smaller liability requires a larger sacrifice of synergies to restore full auditor independence, and vice versa. Accordingly, the adverse impact of restricting NAS is larger (in terms of the synergy loss and its negative effect on audit effort) and thus dominates its positive impact on auditor independence when the auditor's legal liability is smaller. Therefore, when the auditor's legal liability is small, restricting NAS not only reduces audit effort but also impairs the audit quality as well as audit value.

The accounting profession and others opposed to limiting the scope of NAS warn of their negative consequences on audit effort, audit quality, and audit value. Our analysis indicates that restricting NAS has some unintended negative consequences when the auditor's legal liability is sufficiently small. By contrast, when the auditor's legal liability is sufficiently large, restricting NAS not only restores full auditor independence, but also improves audit quality.

Our results thus demonstrate that the effects of the NAS restriction on audit quality and audit value crucially depend on the severity of the auditor's legal liability already in place.

Our research contributes to the literature about restricting the joint provision of NAS to audit clients. Prior research indicates that the provision of NAS creates economic bonds that weaken an auditor's independence, leading to a deterioration of audit quality (DeAngelo 1981; Simunic 1984; Beck et al. 1988). Ewert (2004) calls for theoretical research that studies interdependencies between audit and NAS with regard to possible consequences for auditor independence. Several subsequent studies respond to this call and specifically examine the impacts of restricting NAS on auditors' effort and report decisions. Dopuch et al. (2004) show that restricting NAS may directly broaden or narrow the range of misstatement risk that induces auditor independence and will indirectly broaden the range through increasing audit effort. Our paper supplements theirs by examining whether and under what conditions restricting NAS improves audit quality. Lu and Sapra (2009) investigate a key feature of the violation of auditor independence: biasing opinion in favor of the audit client's interest. Our paper differs from Lu and Sapra (2009) by focusing on another crucial aspect of the violation of auditor independence: suppressing unfavorable audit evidence, which is detrimental per se to audit quality. In Mahieux (2022), successfully detecting fraud gives the auditor some potential power to bargain with the manager to gain the NAS (synergy is assumed away); this gives the auditor extra incentives to exert effort. By contrast, in our model, audit effort in detecting fraud increases the auditor's knowledge about the firm and thus increases the probability of creating value for the firm due to the existence of synergies between audit and NAS; this increases the marginal benefit of audit effort. Our paper also differs from these three papers in that our model endogenizes the synergies between audit and NAS. Other papers relevant to restricting NAS either ignore the auditor's reporting decision (Wu 2006; Friedman and Mahieux 2021) or do not consider the auditor's effort decision (Kornish and Levine 2004).

Our paper also belongs to a burgeoning literature on the economic consequences of heightened audit regulation. Prior studies have examined the effects of increased auditor liability (Dye 1993; Chan and Pae 1998; Hillegeist 1999; Pae and Yoo 2001; Chan and Wong 2002; Laux and Newman 2010; Deng et al. 2012), internal control reporting regulation (Patterson and Smith 2007; Chan 2018), tightening auditing standards (Gao and Zhang 2019; Ye and Simunic 2022), and enhancing the communicative value of audit reports (Chen et al. 2019; Chan and Liu 2023). Our paper complements this literature by studying the effects of restricting NAS and show how the effects depend on auditor liability.

The rest of the paper is organized as follows. Section II lays out the model. Section III characterizes the equilibrium. Section IV analyzes the impact of restricting NAS. We discuss the regulatory implications in Section V. The final section concludes. All proofs are provided in the Appendix.

II. THE MODEL

Consider an economy that spans one period with four dates: 0, 1, 2, and 3, and two groups of risk-neutral players: prospective investors of a representative firm and an auditor.[4] Figure 1 depicts the sequence of events and Table 1 lists the notation in the model. Details of the model are provided below.

[Insert Figure 1 and Table 1 about here]

Investment Opportunities — Initial Investment Project and Expansion Option

At date 0, the firm owns an investment project that requires a capital outlay I from the investors. Once invested, the project's future payoff depends on the value of firm's beginning-of-period assets, which we refer to as its type. The firm's type is ex ante unknown to the investors or the auditor, but it is common knowledge that the firm's type is either good (G) with probability ϕ or bad (B) with the complementary probability, where $\phi \in (0, 1)$. If the firm's type is good, the project will succeed and generate a positive cash flow of $R > I$ at date 2. If the firm's type is bad, the project will succeed and generate R with probability $p \in (0, 1)$ or zero otherwise.

An expansion option will also be available to the firm at date 3 if, and only if, the project has succeeded.[5] With consulting inputs from a business consultant (who could be the incumbent auditor of the firm), the expansion option will provide the firm an additional nonnegative random cash flow of $K \in [\underline{K}, \bar{K}]$. We assume that

$$[\phi + (1-\phi)p]R > I > p(R + \bar{K}). \tag{1}$$

[4] We assume risk neutrality to remove concerns about risk sharing and thus focus on incentive problems alone.

[5] We implicitly assume that the expansion option is available to the firm due to learning by doing from implementing the project.

The first inequality of Condition (1) ensures that the investors should fund the project given their common prior expectation of the firm's type.[6] However, as implied by the second inequality, they should not fund the project if they know that the firm's type is bad, even if they can receive the maximum value of the expansion option.[7]

Audit Services

We assume that the firm always claims that its own type is good, and therefore the unaudited "favorable" financial report is not credible without any attestation by an auditor.[8] To enhance credibility and therefore facilitate the investors' decision, the firm might hire an auditor to ascertain the firm's type. If hired, the auditor receives a flat audit fee, F, to compensate for the audit services.[9] We assume that the auditor possesses a noisy audit technology, based on which the conditional probabilities of audit evidence, g or b, being privately observed by the auditor, are characterized by

$$\Pr(g \mid G, a) = 1 \qquad \text{and} \qquad \Pr(b \mid B, a) = a, \tag{2}$$

[6] Our assumption that ex ante the project (without the expansion option) has a positive net present value (NPV) follows the prior literature (Chan and Pae 1998; Hillegeist 1999; Pae and Yoo 2001; Gao and Zhang 2019) and comports with an economy in which the majority of firms have enormous growth potential. Investors in such an economy are willing to invest in risky investments even without acquiring new information, owing to the large first-mover advantages and handsome return from investing in one of the successful investment projects. The value of the audit is derived from screening out the bad projects. As will be shown clearly in our proofs, the results of our model do not depend on the sign of $[\phi + (1-\phi)p]R - I$. Our main results would therefore remain qualitatively unchanged under an alternative assumption that the project has a negative NPV, wherein the value of the audit is derived from identifying the good projects. Such an alternative assumption, however, requires us to impose restrictions on our model parameters to ensure that the audit opinion is useful for the investment decision in equilibrium; i.e., the investment will be undertaken if and only if the auditor issues an unmodified opinion.

[7] Note that if K is sufficiently large, the investors might not care whether the firm's type is good or bad, thereby eliminating any assurance value of an audit. The second inequality of Condition (1) ensures that this scenario would not occur.

[8] Like ours, most prior studies in the auditing literature assume an exogenous, non-strategic financial report and thus essentially focus on a misreporting equilibrium (Dye 1993, 1995; Chan and Pae 1998; Pae and Yoo 2001; Chan and Wong 2002; Liu and Simunic 2005; Lu and Sapra 2009; Laux and Newman 2010). Exceptions including Deng et al. (2012) and Chan (2016), which consider a strategic reporting decision made by owners/managers. The two papers show that the qualitative results of their papers remain unchanged irrespective of whether the owner/manager employs a fixed or mixed reporting strategy, suggesting that the exogenous, non-strategic financial reporting assumption might actually be innocuous.

[9] The audit fee is independent of the unobservable audit effort. Moreover, to be consistent with the Code of Ethics' ban on contingent fees (rule 302), the audit fee is independent of the audit report.

where we interpret $a \in [0, 1]$ as either the level of effort or quality of auditing technology privately chosen by the auditor. In words, Expression (2) states that for any given audit effort *a*, the auditor never makes a type I error by incorrectly identifying a good-type firm as bad, but may make a type II error, with probability $1 - a$, by incorrectly identifying a bad-type firm as good.[10] The audit effort entails a cost of $\frac{1}{2}ca^2$ on the auditor, where $c > 0$ is a sufficiently large parameter to ensure that the optimal audit effort never exceeds one.

At the end of the audit process, the auditor forms an opinion on the firm's true type. The auditor either agrees with the firm's original favorable report and issues an unmodified opinion (indicated by $\hat{G}$) or disagrees with the firm's report and issues a modified opinion (indicated by $\hat{B}$). The auditor does not necessarily report faithfully in accord with the audit evidence s/he gathered. Specifically, we assume that the auditor reports faithfully if the favorable audit evidence *g* is found (as s/he cannot fabricate evidence to report otherwise), but may choose to suppress the unfavorable audit evidence *b*; i.e.,

$$\Pr(\hat{G} \mid g) = 1 \qquad \text{and} \qquad \Pr(\hat{B} \mid b) \in [0, 1], \tag{3}$$

where $\Pr(\hat{B} \mid b)$ is endogenously determined.[11]

Investment and Outcome of the Project

After the auditor publicly issues an audit report, the investors update their beliefs about the firm's type and decide whether to finance the project. The investors are willing to finance the project if they can expect to break even, and their expectation of the firm's future payoff depends on how they perceive the firm's type.

[10] This assumption is consistent with the notion of discovery sampling that the auditor can report a bad-type firm only when the evidence indicates that the firm's type is bad. In practice, if the auditor falsely accuses the firm of a material error, the resulting denial would trigger additional audit work to rectify the accusation (Arens and Loebbecke 1981, p.136). That is, type I errors would typically be corrected before the auditor issued the audit report.

[11] We assume that if the auditor is indifferent between reporting faithfully and suppressing the unfavorable audit evidence, s/he reports faithfully.

When interpreting the audit opinion, investors must take into account the report incentives of the auditor as well as the imperfections in the audit technology. Given Conditions (1)-(3), the firm's type must be bad upon observing audit report $\hat{B}$ and the investors will never provide fund *I* for the bad-type firm. The game will then end. By contrast, upon observing audit report $\hat{G}$, the investors make the required investment *I*.

At date 2, the cash flow of the project is realized, and the bad-type firm's true type becomes publicly known if the project fails. An audit failure occurs when the auditor has opined that the firm is a good type but the firm's true type is revealed to be bad following the project failure (i.e., a type II error). In the event of an audit failure, the auditor pays a liability payment *L* to the investors.[12]

NAS and Synergies

At date 3, if the project has succeeded, the expansion option arrives. The firm needs consulting inputs from its incumbent auditor or an external business consultant to exercise the expansion option.

We assume for simplicity that if the firm hires an external business consultant, the distribution of K is degenerating to its lowest bound $\underline{K}$. Accordingly, we interpret a value of $K > \underline{K}$ as a synergy arising from the relationship between the firm and its incumbent auditor through the joint provision of audit and NAS and cost savings in switching and adapting to a new business partner (Antle and Demski 1991; Antle et al. 2006).[13]

[12] For simplicity, we take the auditor liability payment *L* to be fixed, independent of the investors' financial loss from relying on an unfaithful audit report.

[13] The incumbent auditor would be the most efficient service provider for many consulting services (e.g., systems design, compliance-related services, and assurance-related services) because the auditor has acquired knowledge of the client's business models and growth opportunities through the audit process. This client-specific knowledge enables the auditor to better identify consulting opportunities and to lower the costs of marketing and providing NAS to his/her client. We also implicitly assume that these services are contractible.

We model the client-specific knowledge that contributes to the synergy as the knowledge spillovers from audit to nonaudit services. Thus, we assume that the distribution of K is positively affected by the amount of the audit effort exerted in the audit process.[14] Let ωa, where $\omega \in (0,\ 1)$, be the probability that the synergy is uniformly distributed in the interval $(\underline{K},\ \bar{K}]$ when the audit effort is a. With the remaining probability, i.e., $1-\omega a$, the synergy is equal to $\underline{K}$. The value of ω captures the effectiveness of the audit effort in generating a higher synergy from the joint provision of audit and NAS. Let $\Omega(K\,|\,a) \equiv 1-\omega a(1-\frac{K}{\bar{K}}) \in [0,\ 1]$ be the associated cumulative distribution function of $K \in [\underline{K},\ \bar{K}]$ conditional on a.[15] The cumulative distribution function is continuously differentiable over $K \in [\underline{K},\ \bar{K}]$ and over $a \in [0,\ 1]$, and satisfies the first order stochastic dominance property: $\partial\Omega(K\,|\,a)/\partial a = -\omega(1-\frac{K}{\bar{K}}) < 0$ for all $K \in (\underline{K},\ \bar{K})$. Because of the existence of synergy, the firm will hire its incumbent auditor for consulting when exercising the expansion option.

Cap on NAS, Auditor Independence, and Audit Quality

The actual amount of the benefit from exercising the expansion option, however, is constrained not only by production technology, but also by the regulation governing the scope of NAS that the auditor can perform along with the audit services provided to the same client. In this paper, we model the impact of an NAS restriction by assuming a cap on the allowed benefit of consulting, which is increasing in the scope of NAS provided to the audit client. Let

[14] The idea is that an increase in audit effort enables the auditor to acquire more knowledge of the nature of the client's business and industry, accounting policies, levels of control, and other risks that might cause accounting problems. Acquiring this knowledge will not only enhance the audit effectiveness and efficiency, but also help generate and perform more consulting tasks.

[15] Arguably, the good-type firm is more likely to provide better consulting opportunities to the auditor. In an extreme case where the expansion option is present only when the firm type is good, the auditor would have no incentive to suppress the unfavorable evidence *b* once s/he identified the bad firm type. In this case, the issue of misreporting audit evidence would be moot. To emphasize the impact of NAS on auditor independence and simplify the analysis, we assume that the distribution of *K* is independent of the firm's type.

$Y \leq \bar{K}$ be the maximum amount of K under the NAS restriction. Thus, the actual amount of K under the regulation is given by $X(Y) = \min\{K, Y\}$. In other words, a restriction on NAS reduces the maximum potential economic benefits of the joint provision of audit and NAS from $\bar{K}$ to Y.[16] Thus, we can study the consequence of a restriction on NAS by analyzing the impact of changing Y and, specifically, use the case in which $Y = \bar{K}$ to capture the pre-regulation scenario.

In our model, the auditor forfeits the opportunity to earn consulting fees whenever s/he issues a bad report. Thus, by design, the consulting fees in our model are de facto contingent.[17] Moreover, $X(Y)$ is a quasi-rent to the incumbent auditor. For simplicity, we assume that the incumbent auditor will keep a fraction τ of $X(Y)$ as his/her consulting fee when the project succeeds, where $\tau \in (0, 1]$ measures the relative bargaining strength of the auditor and is assumed to be exogenously determined outside the model and not to vary across different regulatory regimes.[18]

To address the popular claim that auditors might compromise their independence in reporting when they are free to provide any NAS to their audit clients, we further assume that

[16] We model the impact of a NAS restriction by a cap on, rather than a proportional reduction of, the synergy K to better capture certain salient features of such regulation in realty. For example, Section 201 of SOX in the U.S. prohibits auditors from providing their audit clients with human resources. The synergy benefit from jointly providing human resources is completely, not proportionally, lost under Section 201. On the other hand, auditors can still provide, for example, tax service as long as audit committees have reviewed and approved them. There is no restriction imposed by Section 201 on the proportion of tax services that the auditor can provide to his/her audit client. The European Union introduced a similar legislation in 2016, which has also put a cap on permissible NAS of 70% of the average of the audit fees paid in the last three consecutive financial years. The idea behind our approach to model the impact of a NAS restriction is that, even if many consulting opportunities might emerge from auditing, the maximum synergy benefit cannot exceed a threshold $Y < \bar{K}$ because legislation restricts the scope of NAS jointly provided to the same audit client.

[17] The Code of Ethics' ban on contingent fees (rule 302) does not prohibit these types of contingent fees from sources other than auditing.

[18] The auditor has no strong incentive to suppress the unfavorable audit evidence when $\tau = 0$. In the other extreme, the client might still have incentive to hire the incumbent auditor for the consulting even if $\tau = 1$ because the auditor-provided NAS enhances the ex ante incentive effect on the auditor's effort and thus results in a higher assurance value of the audit.

the auditor privately observes a perfect signal about the realization of K after observing audit evidence, g or b, but prior to issuing an audit opinion, $\hat{G}$ or $\hat{B}$.[19]

When making the report decision, the auditor compares the expected liability loss with the consulting profit based on the audit evidence s/he collected in the audit process. Specifically, when choosing whether to issue a bad report after observing the unfavorable audit evidence b and the realization of $X(Y)$, the auditor weights the expected profit from providing NAS to the same client (i.e., $p\tau X(Y)$) against the expected liability loss (i.e., $(1-p)L$). To ensure that the joint provision of audit and NAS might pose threats to auditor independence as the auditor might issue an unwarranted audit opinion for fear of losing lucrative consulting fees, we assume that [20]

$$p\tau\overline{K} > (1-p)L > p\tau\underline{K}. \tag{4}$$

In other words, we restrict our attention to $L \in (\underline{L}, \overline{L})$, where $\underline{L} \equiv p\tau\underline{K}/(1-p)$ and $\overline{L} \equiv p\tau\overline{K}/(1-p)$. Accordingly, for any given Y, the auditor will suppress the unfavorable audit evidence b if $p\tau X(Y) > (1-p)L$, and report faithfully otherwise.[21] That is, there exists a unique threshold

$$\Delta \equiv (1-p)L/p\tau, \tag{5}$$

[19] This assumption is consistent with the view that auditors can learn how to improve their clients' operations in the course of doing audit.

[20] If $p\tau\overline{K} \le (1-p)L$, the auditor's future payoff from consulting is always less than or equal to the expected liability loss, and hence, s/he will always report faithfully even in the temptation of future consulting fees. The role of a NAS restriction on the auditor's reporting behavior then would be moot. Regarding the second inequality, if $p\tau\underline{K} \ge (1-p)L$, the auditor will never report faithfully. Then, a regulator who is concerned about auditor independence should have banned rather than just restricted the provision of NAS. However, a complete ban of NAS provision is not consistent with what we observe in reality. Thus, we believe that Condition (4) is descriptive when a NAS restriction, such as Section 201 of SOX, was enacted as it describes an (unregulated) economy in which the auditor either faithfully or unfaithfully issue an opinion depending on the tradeoff between the expected liability loss and the expected consulting profit.

[21] DeFond et al. (2002) underline that auditors are willing to sacrifice their independence if reputation and litigation costs associated with audit failures are smaller than the economic rents from the provision of NAS to the same client. See also Causholli et al. (2014) for empirical evidence that supports a positive association between selling future NAS and impairment of auditor independence.

such that the auditor will suppress the unfavorable audit evidence b if, and only if, $X(Y) > \Delta$, and will report faithfully otherwise. The second inequality in Condition (4) ensures that there exists a non-empty segment of K wherein the auditor will report faithfully even if there is no NAS restriction, i.e., when $Y = \bar{K}$. In the sequel, we normalize $\underline{K}$ to zero.

The cap on NAS (i.e., Y) influences the auditor's decision problem and hence the game tree. Figure 2 depicts the game tree when $Y > \Delta$.[22]

[Insert Figure 2 about here]

To ensure that the auditor in our model, despite being opportunistic, behaves "nicely," we make the following assumptions concerning parameter values:

$$c > \underline{c} \equiv (1-\phi)[I - p(R + \tfrac{1}{2}\omega\bar{K})] + \tfrac{1}{2}\phi\omega\bar{K}, \qquad (6)$$

$$\phi > \underline{\phi} \equiv \tfrac{p}{1+p} \in (0,\ \tfrac{1}{2}). \qquad (7)$$

Condition (6) ensures that the optimal audit effort is interior. Condition (7) is equivalent to $\phi > (1-\phi)p$ and means that a good-type firm has sufficiently higher value than a bad-type firm.

III. THE EQUILIBRIUM

Before proceeding to solve the model by backward induction, we first highlight several remarks regarding the policy variable Y and introduce the concepts of auditor independence and audit quality in the context of our model. First, the threshold Δ as defined in Equation (5) is independent of Y. Second, when $Y \le \Delta$, a restriction on NAS is effective in restoring full auditor independence. Third, if $Y > \Delta$, the auditor issues an unfaithful audit report whenever $K > \Delta$, despite the regulation. The basic tension in the model is that while a restriction on NAS

[22] The game tree when $Y \le \Delta$ is similar except that the auditor always faithfully issues a modified opinion upon observing an unfavorable audit evidence.

might restore full auditor independence, it also destroys synergies by reducing the expected value of the synergy arising from the joint provision of audit and NAS. Accordingly, auditor independence is the probability that the auditor reports $\hat{B}$ conditional on observing *b*, and audit quality is the joint probability that the auditor correctly identifies and reports the bad-type firm as bad, defined formally below.[23]

Definition 1: *Given any pair of a and Y, auditor independence is denoted by* $AI(a, Y)$ *and equals*

$$AI(a, Y) \equiv \Pr\left(\hat{B} \middle| b, Y\right) = \begin{cases} 1 & \text{if } Y \leq \Delta, \\ \Omega(\Delta|a) & \text{if } Y > \Delta. \end{cases}$$

Audit quality is denoted by $AQ(a, Y)$ *and equals*

$$AQ(a, Y) \equiv \Pr(b|B, a) \Pr\left(\hat{B} \middle| b, Y\right) = a(Y)AI(a, Y) = \begin{cases} a(Y) & \text{if } Y \leq \Delta, \\ a(Y)\Omega(\Delta|a) & \text{if } Y > \Delta. \end{cases}$$

The auditor is *fully independent* if his/her opinion is strictly consistent with his/her privately observed audit evidence; i.e., $AI(a, Y) = 1$. However, even if the auditor is fully independent in reporting, audit quality could still be low because audit effort could be low. Lemma 1 below provides properties of auditor independence and audit quality.

Lemma 1: *Auditor independence* $AI(\mathrm{a}, \mathrm{Y})$ *and audit quality* $AQ(\mathrm{a}, \mathrm{Y})$ *have the following properties.*

i. Given any audit effort $a > 0$, auditor independence and audit quality display a discrete, negative jump at $Y = \Delta$.

[23] Our definitions of auditor independence and audit quality are consistent with those commonly used in the auditing literature (see DeAngelo 1981, p. 186). Unlike a large strand of the theoretical auditing literature that implicitly assumes the choice of audit effort equal to audit quality, audit quality is different from audit effort in our model because audit quality is jointly affected by the auditor's effort and report decisions.

ii. For $Y \leq \Delta$, restricting NAS effectively induces full auditor independence regardless of audit effort, whereas audit quality is solely determined and linearly increasing in audit effort.

iii. For $Y > \Delta$, restricting NAS does not affect auditor independence holding audit effort fixed and thus cannot induce full auditor independence; in addition, auditor independence strictly decreases in audit effort, whereas audit quality is increasing in audit effort.

iv. For $Y > \Delta$ and given any audit effort $a > 0$, auditor independence and audit quality strictly increase in auditor liability L.

Properties (i) and (ii) follow immediately from our model assumptions. The intuition behind Property (iii) is as follows. An increase in audit effort increases the probability of realizing a larger synergy and, when it happens, the synergy induces the auditor to compromise. However, an increase in audit effort also increases the probability of detecting a misstatement, and therefore its impact on audit quality is potentially ambiguous. We show that the latter effect always dominates the former effect and therefore the marginal probability of issuing a modified report is increasing in audit effort even for an "unfaithful" auditor. For Property (iv), larger liability payments strictly increase the threshold Δ, which in turn leads to a larger $AI(a, Y > \Delta)$, i.e., $d\Delta/dL = (1-p)/p\tau > 0$ and $\partial AI(a, Y > \Delta)/\partial\Delta = \partial\Omega(\Delta|a)/\partial\Delta = \omega a/\bar{K} > 0$. Also observe that, in the limit when $L = \Delta = 0$, auditor independence and audit quality have a lower bound of $AI(a, Y) = \Omega(\Delta|a) = 1 - \omega a$ and $AQ(a, Y) = a\Omega(\Delta|a) = a(1 - \omega a)$, respectively. [24]

[24] The lower bound of auditor independence is due to the stochastic nature of synergies. Although the auditor's effort stochastically increases the synergies, there is always a positive probability that the auditor's effort does not generate any synergies.

Investors' Decision

After observing the audit opinion, the investors decide whether to make the required investment to undertake the project. Although the investors do not know the realization of *K* when a good audit report is observed, they know that the auditor will optimally suppress the unfavorable audit evidence *b* if, and only if, $X(Y) > \Delta$.The investors know that the auditor issues a good audit report in one of three different states. First, the auditor accurately issues a faithful, favorable audit report to the good-type firm with probability $Pr(\hat{G}|g)\,Pr(g|G,a')\,Pr(G) = \phi$, where a' denote investors' conjectured audit effort (which must be self-fulfilling in a rational expectation equilibrium). Second, the auditor mistakenly issues a faithful, favorable audit report to the bad-type firm with probability $\Pr(\hat{G} \mid g, Y)\Pr(g \mid B, a')\Pr(B) = (1-a')(1-\phi)$ because s/he fails to identify the bad-type firm. Third, the auditor purposefully issues an unfaithful, favorable audit report to the bad-type firm with probability $\Pr(\hat{G}|b,Y)\Pr(b|B,a')\Pr(B) = \left(1 - AI(a',Y)\right)a'(1-\phi)$, because s/he, albeit successfully identifying the bad-type firm, suppresses the unfavorable evidence to seize the lucrative consulting opportunity.[25] Summing up these probabilities and simplifying terms yields the probability of an unmodified opinion:

$$\Pr(\hat{G}|a',Y) = \phi + (1-\phi)\left(1 - AQ(a',Y)\right), \tag{8}$$

which is decreasing in audit quality.

The investors update the posterior probability that a firm with a favorable audit report is actually of good type to

$$\Pr(G|\hat{G}, a', Y) = \frac{\Pr(G,\hat{G}|a')}{\Pr(\hat{G}|a',Y)} = \frac{\phi}{\phi + (1-\phi)\left(1 - AQ(a',Y)\right)}. \tag{9}$$

[25] Note that the third state never occurs if $Y \leq \Delta$.

Observe from Equation (9) that $\Pr(G|\hat{G}, a', Y) > \phi$. The investors are willing to make the required investment whenever the auditor issues a good audit report. Accordingly, at date 0, investors' unconditional net expected return from investing in the project without the expansion option is given by

$$ER(a', Y) = \left[\phi + (1-\phi)\left(1 - AQ(a', Y)\right)p\right]R - \left[\phi + (1-\phi)\left(1 - AQ(a', Y)\right)\right]I. \quad (10)$$

Moreover, if the auditor issues report $\hat{G}$ but the project fails, investors will recoup L through suing the auditor. At date 0, the auditor has not exerted any effort yet and hence does not observe K. Thus, the investors' unconditional expected payoff from the litigation against the auditor equals the auditor's expected litigation payment $EL(a', Y)$, which will be provided and discussed in detail shortly.

The investors can also earn $(1-\tau)$ proportion of the synergy. At date 0, the total unconditional expected synergy generated by the auditor is given by

$$ES(a', Y) = [\phi + (1-\phi)(1-a')p]E[X(Y)|a'] + 1_{Y>\Delta}(1-\phi)a'p\{E[X(Y)|a'] - AI(a', Y > \Delta)E[K|K \leq \Delta; a']\}, \quad (11)$$

where

$$E[X(Y)\,|\,a'] = \int_0^Y K d\Omega(K\,|\,a') + \int_Y^{\bar{K}} Y d\Omega(K\,|\,a') = \omega a' Y(1 - \tfrac{Y}{2\bar{K}})$$

is the expected value of the legally allowed benefit from the joint provision of audit and NAS to the firm, and

$$E[K\,|\,K \leq \Delta; a'] = \Delta - \frac{\int_0^\Delta \Omega(K\,|\,a')dK}{\Omega(\Delta\,|\,a')} = \frac{\omega a' \frac{\Delta^2}{2\bar{K}}}{1 - \omega a'(1 - \frac{\Delta}{\bar{K}})}$$

is the conditional mean of K given that $K \leq \Delta$, and $1_{Y>\Delta}$ is an indicator function that satisfies

$$1_{Y>\Delta} = \begin{cases} 1 & \text{if } Y > \Delta, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the auditor makes no mistake for the good-type firm and therefore will always provide consulting businesses to the good-type firm. By contrast, the auditor loses the consulting opportunity whenever s/he issues a modified audit opinion to the bad-type firm. Accordingly, if the auditor is fully independent, the synergy arises only if the auditor obtains the favorable audit evidence g and the project succeeds, so the expected value of K is given by the first term of Expression (11). In the regime $Y > \Delta$, the auditor is not fully independent and purposefully suppresses unfavorable audit evidence b (if it exists) when $K > \Delta$, thereby increasing the expected value of NAS by the second term of Expression (11).

Adding all these components together, the investors' unconditional expected payoff at date 0 is given by

$$U\left(a',Y\right) = ER\left(a',Y\right) + EL\left(a',Y\right) + (1-\tau)ES\left(a',Y\right) - F. \tag{12}$$

If no auditor were hired, the investors' decision will be based on the common prior beliefs about the firm type, and the expansion option adds no value to the firm. The investors' expected payoff in the case of no audit is then given by

$$[\phi + (1-\phi)p]R - I. \tag{13}$$

Subtracting Expression (13) from Expression (12) yields the value of the audit to the firm (represented by the increase in the investors' expected payoff due to the audit), which is given by

$$A\left(a',Y\right) = U\left(a',Y\right) - \{[\phi + (1-\phi)p]R - I\}. \tag{14}$$

Auditor's Effort and Report Decisions

The auditor incurs legal liability if (i) the firm is a bad-type and fails and (ii) the auditor issues an unmodified audit opinion due to either failing to detect the bad-type firm or suppressing the unfavorable evidence b. The corresponding probability of the latter is given by $1 - AQ(a,Y)$. Therefore, the auditor's unconditional expected legal liability before exerting effort is given by

$$EL(a,Y) = (1-\phi)\big(1 - AQ(a,Y)\big)(1-p)L. \tag{15}$$

If the project succeeds, the auditor will provide NAS and keep τ proportion of $X(Y)$ as his/her consulting fee. Prior to collecting the audit evidence, K is unknown to the auditor. Thus, given any a, at date 0 the auditor has the same total unconditional expected synergy as the investors': $ES^*(a,Y)$. Then, the auditor's unconditional expected profit when s/he is hired is given by

$$\Pi(a,Y) = F - \frac{1}{2}ca^2 - EL(a,Y) + \tau ES(a,Y). \quad (16)$$

We first examine a hypothetical setting (which we refer to as the first-best case) where the auditor commits to report faithfully and expends effort to maximize the (social) value of audit (as dictated by a "social planner"). Using Equations (14) and (16), the (social) value of the audit, as defined by the sum of the value of the audit to the firm and the auditor's expected profit, is given by[26]

$$V(a,Y) = A(a,Y) + \Pi(a,Y) = (1-\phi)AQ(a,Y)(I-pR) + ES(a,Y) - \frac{1}{2}ca^2. \quad (17)$$

Inspection of Equation (17) reveals that $V(a,Y)$ consists of three components: The first term is the saving of the investment deadweight loss, which depends on the audit quality and reflects the assurance value of the audit. The second term is the expected additional benefit (synergy) from the joint provision of audit and NAS. The last term is the audit cost. Neither the share rule τ nor the liability payment L appear in $V(a,Y)$ because these represent transfers among players and do not affect aggregate outcomes. [27]

In the first-best case, there is no concern of auditor independence, i.e., $AI(a,Y) = \Pr(\hat{B}|b,Y) = 1$ for all Y, and thus there is no need to impose any restriction on NAS; i.e., $Y = \overline{K}$. Accordingly, $AQ(a,Y) = a$ and $ES(a,Y) = \frac{1}{2}[\phi + (1-\phi)(1-a)p]\omega a\overline{K}$ in Equation

[26] Note that the difference between audit value V and social welfare (i.e., the sum of the expected payoffs of all players in the model) is a constant term $[\phi+(1-\phi)p]R-I$. Therefore, our analysis of audit value is equivalent to social welfare analysis.

[27] We do not consider litigation frictions; thus, the auditor's legal liability loss is a transfer payment.

(17). The social planner's problem is to choose a to maximize $V(a,\bar{K})$. Using the superscript "FB" to denote the first-best equilibrium, the solution to this maximization problem is a^{FB}, which is given by

$$a^{FB}=\frac{(1-\phi)(I-pR)+\frac{1}{2}[\phi+(1-\phi)p]\omega\bar{K}}{c+(1-\phi)p\omega\bar{K}}<1, \tag{18}$$

where the inequality follows from Condition (6) that $c>\underline{c}\equiv(1-\phi)[I-p(R+\frac{1}{2}\omega\bar{K})]+\frac{1}{2}\phi\omega\bar{K}$.

We now return to the original setting in which the auditor is opportunistic with respect to both his/her effort and report decisions. The auditor's decision problem at date 0 is to choose an audit effort a to maximize Equation (16). Using the superscript "*" to denote the (second-best) equilibrium, the interior solution to this decision problem, $a^*(Y)$, is given by[28]

$$a^*(Y)=\frac{(1-\phi)p\Delta+[\phi+(1-\phi)p]\omega Y\left(1-\frac{Y}{2\bar{K}}\right)}{\frac{c}{\tau}+2(1-\phi)p\omega\min\left\{\Delta\left(1-\frac{\Delta}{2\bar{K}}\right),Y\left(1-\frac{Y}{2\bar{K}}\right)\right\}}. \tag{19}$$

IV. THE EFFECTS OF NAS RESTRICTION

The Effect of NAS Restriction on Audit Effort and Auditor Independence

We now study the economics consequences of restricting NAS. Inspection of Equation (19) reveals that the (second-best) equilibrium audit effort largely depends on the NAS allowed to be jointly provided with audit. Proposition 1 below states formally the impact.

Proposition 1: *The equilibrium audit effort $a^*(Y)$ is always below a^{FB} and increasing in Y.*

The intuition of Proposition 1 is that the auditor faces the same marginal cost but lower marginal benefits than the social planner. Thus, the second-best solution entails a lower audit effort than the first-best. Proposition 1 implies that the social problem, when neither the audit

[28] Proposition 1 below shows that $a^*(Y)<a^{FB}$ and therefore $a^*(Y)$ must not exceed 1.

effort is observable nor the auditor is fully independent, is one of under-provision in audit effort. The question is then whether restricting NAS can alleviate this under-provision problem in audit effort.

[Insert Figure 3 about here]

Proposition 1 continues to state that restricting NAS results in a decrease in audit effort. Figure 3 depicts this result. A restriction on NAS caps the auditor's maximum consulting fee to τY with $Y < \bar{K}$. Given that audit effort increases the probability of obtaining $K > Y$, decreasing Y has a *direct* effect that reduces the auditor's incentive to work hard by decreasing the marginal benefit of audit effort. In addition, as Y decreases from $\Delta^{+} \equiv \Delta + \varepsilon$ to $\Delta^{-} \equiv \Delta - \varepsilon$, where ε is an arbitrarily small number, the regulation's effect in inducing the auditor to be fully independent changes from ineffective to effective. Accordingly, a decrease in Y has the following two additional effects.

First, working harder could lower the probability of an audit failure. When $Y = \Delta^{+}$ the auditor has less incentive to work hard to detect fraud than when $Y = \Delta^{-}$, as s/he may not use the obtained evidence anyway when $Y = \Delta^{+}$. Hence, a decrease in Y from Δ^{+} to Δ^{-} has an *indirect* effect that induces the auditor to work harder, as audit effort is more effective in avoiding liability losses. Second, working harder could also risk losing the consulting profit once the auditor identifies the bad-type firm in the audit process. This is more pronounced when $Y = \Delta^{-}$ than when $Y = \Delta^{+}$, since the auditor is induced to report faithfully when $Y = \Delta^{-}$. Consequently, a decrease in Y from Δ^{+} to Δ^{-} has another opposite *indirect* effect that induces the auditor to choose a lower level of audit effort to lower the risk of losing the consulting profit. In the end, these two opposite *indirect* effects balance out and have no net impact on the equilibrium audit effort. Hence, the overall effect of NAS restriction on audit effort is determined by the remaining *direct* negative effect, suggesting that the auditor chooses a lower audit effort with the NAS restriction. [29]

Besides, when $Y < \Delta$ the auditor gives up the chance to earn consulting fee if seeing *b*; therefore, the cap on NAS has less impact on the auditor's expected consulting fee when $Y < \Delta$ than when $Y > \Delta$. Thus, audit effort increases faster at $Y = \Delta^{+}$ than at $Y = \Delta^{-}$.

Since $\Omega(\Delta|a)$ is decreasing in audit effort a, the following corollary regarding the effect of increasing *Y* on auditor independence $AI(a^{*}(Y), Y)$ is immediate.

Corollary 1: *Auditor independence* $AI(a^{*}(Y), \mathrm{Y})$ *is independent of Y when* $Y \leq \Delta$ *and decreasing in Y when* $Y > \Delta$.

Figure 3 depicts the above results. Opponents of NAS restriction argue strongly that restricting NAS may damage the synergy between audit and NAS, thereby reducing the auditor's incentive to exert effort. Proposition 1 lends support to this concern. By contrast, proponents of the NAS restriction would argue that an increase in *Y* leads to a decline in audit independence *AI*, and given any effort *a*, a decline in audit quality *AQ* as well. Corollary 1 lends support to this concern. Regulators, investors, and accounting practitioners are more likely concerned with the economic effects of changing *Y* on audit quality and the (social) value of the audit, which we are going to examine next.

The Effect of NAS Restriction on Audit Quality

Proposition 2 below elaborates the impact of changing *Y* on audit quality.

[29] Notice that the two opposite *indirect* effects do not balance out by coincidence: Δ is the unique threshold of *K* at which the auditor's expected legal liability from suppressing unfavorable audit evidence equals his/her expected consulting fee from doing so, and, thus, is also the largest *Y* that ensures that the auditor is fully independent.

Proposition 2: *When the NAS allowed to be jointly provided with audit (i.e., Y) increases in the range* $[0, \overline{K}]$, *audit quality, i.e.,* $AQ(a^*(Y), Y)$, *increases for* $Y \in [0, \Delta)$, *jumps down at* $Y = \Delta$, *and then increases for* $Y \in (\Delta, \overline{K}]$. *In addition, there exists a unique threshold,* $\hat{L} \in (0, \overline{L})$, *such that* $AQ(a^*(\Delta), \Delta) < AQ(a^*(\overline{K}), \overline{K})$ *if, and only if,* $L < \hat{L}$.

We graphically illustrate the above results in Figure 4. The graph of $AQ(a^*(Y), Y)$ breaks at $Y = \Delta$. The downward jump is due to the drastic change in the auditor's independence: a restriction of NAS is effective (ineffective) in inducing full auditor independence if $Y \leq \Delta$ ($Y > \Delta$). Holding the auditor's reporting strategy fixed (i.e., given that Y changes within one of the two regions $[0, \Delta)$ and $(\Delta, \overline{K}]$), $AQ(a^*(Y), Y)$ increases in Y (as reflected by the upward slope of each curve). The reason is that an increase in Y enhances the potential synergy between audit and NAS and thus induces more audit effort, as stated in Proposition 1.

[Insert Figure 4 about here]

As the whole graph comprises two upward-sloping curves, a comparison between the two boundaries $AQ(a^*(\Delta), \Delta)$ and $AQ(a^*(\overline{K}), \overline{K})$, which represent the highest possible audit quality that can be obtained with and without the NAS restriction, respectively, is necessary to analyze the effect of restricting NAS. Proposition 2 identifies a necessary and sufficient condition under which the "synergy-induced-effort" effect overwhelms the "reporting-strategy-shifting" effect on audit quality. Specifically, a restriction of NAS that can successfully restore full auditor independence undermines audit quality if, and only if, the auditor's legal liability is sufficiently small; i.e., $L < \hat{L}$. The rationale is as follows. As discussed, the auditor weights the expected consulting profit (which is derived from the synergy between audit and NAS) against the expected liability loss to determine whether to suppress the unfavorable audit evidence. Thus, the threshold of additional benefit K for the auditor to suppress the unfavorable audit evidence (i.e., $\Delta \equiv (1-p)L / p\tau$) is increasing with the auditor's liability loss L. The smaller the auditor's liability loss, the more synergy that must be sacrificed to restore full auditor independence. Thus, when the auditor's liability loss is sufficiently small, the regulation that effectively restores full auditor independence will impose a significant negative effect on audit effort due to the required tremendous adverse impact on the synergy loss. As such, the negative effect on audit effort of restricting NAS dominates its positive effect on restoring full auditor independence, thereby leading to an overall decrease in audit quality.

The Effect of NAS Restriction on Audit Value

As shown in Equation (17), the value of the audit equals the assurance value of the audit plus the synergy from NAS minus the audit cost. Inspection of Equations (10) and (11) reveals that the assurance value of the audit is increasing in audit quality (i.e., it is increasing in audit effort and is smaller if the auditor suppresses any unfavorable evidence), whereas the expected synergy is increasing in audit effort but is larger if the auditor has an independence issue. Proposition 3 below sheds some light on the impact of changing Y on the value of the audit.

Proposition 3: *When the NAS allowed to be jointly provided with audit (i.e., Y) increases in the range* $[0, \overline{K}]$*, the value of the audit, i.e.,* $V^*(a^*(Y), Y)$*, increases for* $Y \in [0, \Delta)$*, jumps down at* $Y = \Delta$ *, and then increases for* $Y \in (\Delta, \overline{K}]$ *. In addition, a sufficient condition for* $V^*(\, a^*(\Delta), \Delta) < V^*(\, a^*(\overline{K}), \overline{K})$ *is that* $L < \hat{L}$*.*

Figure 5 graphically depicts the results stated in Proposition 3. Similar to the graph of $AQ(a^*(Y), Y)$ as shown in Figure 4, the graph of $V^*(a^*(Y), Y)$ also breaks at $Y = \Delta$ due to the change in the auditor's independence. Recall that Figure 3 shows that audit effort $a^*(Y)$ is continuous in Y, so audit cost remains unchanged when Y increases from Δ^- to Δ^+. When Y

increases from Δ^- to Δ^+, the auditor changes his/her reporting strategy from full auditor independence to suppressing the unfavorable evidence when observing a large K, so the change in Y enhances synergy but sacrifices the assurance value of the audit. The first part of Proposition 3 states that the former synergy effect is dominated by the latter assurance value effect, resulting in a downward jump.

[Insert Figure 5 about here]

Holding the auditor's reporting strategy fixed (i.e., given that Y changes within one of the two regions $[0, \Delta)$ and $(\Delta, \overline{K}]$), the value of the audit $V^*(a, Y)$ increases in Y (as reflected by the upward slope of each curve). The reason is that an increase in Y enhances the synergy and thus induces a higher audit effort, which further improves the assurance value of the audit.

Similar to the analysis of Proposition 2, the comparison between the boundaries $V^*(a^*(\Delta), \Delta)$ and $V^*(a^*(\overline{K}), \overline{K})$, which represent the highest possible audit value that can be obtained with and without the NAS restriction, respectively, becomes necessary when the curve is not continuous. We proceed by seriatim, comparing the three components of the value of the audit. Firstly, restricting NAS from $\overline{K}$ to Δ directly reduces the maximum synergy and thus reduces the value of the audit. Secondly, a decrease in Y reduces audit effort and thus reduces the audit cost, thereby increasing the value of the audit. The first effect always dominates the second effect since the auditor strategically chooses audit effort (see the proof of Proposition 3). The third component is the assurance value of the audit: it is smaller in the $Y = \Delta$ regime than in the $Y = \overline{K}$ regime if $L < \hat{L}$, because under this condition, as stated in Proposition 2, restricting NAS from $\overline{K}$ to Δ reduces audit quality. Accordingly, Proposition 3 states that if $L < \hat{L}$, the negative net effect of the sum of the first two effects and the negative third effect together leads to an overall deterioration of the value of the audit.

V. REGULATORY IMPLICATIONS

Results from Propositions 1 to 3 showed that if the auditor's legal liability is sufficiently small, i.e., $L < \hat{L}$, then audit effort, audit quality, and the value of the audit are maximized by no restriction on NAS. These results suggest that a regulator should not restrict NAS when the auditor's legal liability is sufficiently small. However, when the auditor's legal liability is sufficiently large, i.e., $L > \hat{L}$, a restriction on NAS (with $Y = \Delta$) can be justified on the ground of an improved audit quality. Hence, how litigious is the audit market when the NAS restriction was enacted plays a crucial role in determining the regulatory impacts. Corollary 2 below provides additional insight into which one of these two situations (i.e., $L > \hat{L}$ or $L < \hat{L}$) is likely to prevail when the restriction on NAS was enacted.

Corollary 2: $\hat{L}$ *is decreasing in* ω.

Corollary 2 states that $\hat{L}$ is smaller with a larger ω, implying that holding L fixed, the NAS restriction is more likely to increase audit quality when the joint provision of audit and NAS is more likely to generate synergy. In the U.S., the auditing industry has experienced dramatic changes since the mid-1970s. On the one hand, consulting services have grown rapidly in the accounting profession since the mid-1970s, and by 2000, the then Big Five audit firms had expanded into multidisciplinary professional service firms that earned more than half of their gross fees from NAS. This trend might be demand-driven due to client business evolutions over time that render the incumbent auditor fit for performing certain NAS; i.e., ω becomes larger and thus $\hat{L}$ becomes smaller. On the other hand, the Private Securities Litigation Reform Act of 1995 has replaced joint-and-several liability with proportionate liability, and most of the national audit firms are now limited liability corporations (which means that an auditor's effective legal liability is indeed the lower of either his/her legal penalty or registered wealth); i.e., L becomes smaller. In this context, whether $L < \hat{L}$ or $L > \hat{L}$ is more descriptive when NAS legislation, such as Section 201 of SOX in the U.S., was enacted is an empirical question. If, in fact, $L < \hat{L}$ and there are no other legislations in tandem that imposed greater legal liability to the auditor, then implementing the NAS restriction might have unintended detrimental economic consequences on audit effort, audit quality, and the value of the audit.

Our study also speaks to the NAS legislation in the EU, because auditor liability there is not as severe as in the U.S. For instance, on 6 April 2008, with the aim of promoting audit market competition, the European Commission issued a recommendation encouraging member states to limit auditor liability.[30] Currently, while major differences exist as to the liability regime of the statutory auditor within the EU, some member states (e.g., Belgium, Germany, and Greece) allow statutory caps of a fixed amount of damages in the case of litigation.

VI. CONCLUSION

The supply of NAS by auditors to their audit clients has long remained a contentious issue in the accounting profession. Various regulatory bodies have suggested that such supplies lead to increased bonding between auditors and their clients, thereby adversely affecting auditor independence. Practitioners, on the other hand, have consistently suggested that joint supplies of NAS and auditing lead to efficiency and improved performance in audits and consulting engagements. In the U.S., the concern about auditor independence is so strong that Section 201 of the SOX of 2002 imposes significant restrictions on NAS that an auditor can perform for an audit client. The EU introduced similar restrictions on NAS in 2016. In the U.K., audit client is now limited to services regarding legally required reports and audit-related services (Financial Reporting Council 2019).

[30] This recommendation rests on the premise that limited liability could lower the insurance premiums, lift a major market-entry barrier for smaller auditing firms, and thus promote audit market competition. See "Commission Recommendation Concerning the Limitation of the Civil Liability of Statutory Auditors and Audit Firms (2008/473/EC)", available at https://ec.europa.eu/transparency/regdoc/rep/2/2008/EN/SEC-2008-1974-F1-EN-MAIN-PART-1.PDF

This paper provides an economic framework to decipher the effects of NAS restriction on auditors' effort and report decisions. In our model, auditors are strategic in choosing both effort and audit report, and the joint provision of audit and NAS to the same client creates synergies. We show that although the NAS restriction can be effective in restoring full auditor independence, it creates perverse incentives to exert a lower audit effort and can, thus, decrease audit quality. Specifically, we show that this regulatory change has unintended negative consequences on audit effort, audit quality, and the value of the audit when the auditor's legal liability is sufficiently small.

We have presented a stylized model to deliver the intuition for economic tradeoff between the NAS restriction's ex ante audit effort and *ex post* audit opinion incentives. In particular, the binary structures with respect to both the measurement and reporting of outcomes have dramatically simplified the exposition. While we believe that the key tension in the model remains intact with continuous structures, the use of binary structures has prevented us from studying some interesting issues such as the determination of auditors' optimal materiality thresholds and how these thresholds influence decisions of financial statement users. These interesting issues are a promising venue for future research.

APPENDIX: PROOFS

Proof of Lemma 1

The proofs of parts (i) and (ii) are by inspection. For part (iii), we have $\partial AI(a, Y > \Delta)/\partial a = \partial \Omega(\Delta|a)/\partial a = -\omega(1-\frac{\Delta}{\bar{K}}) < 0$ and $\partial AQ(a, Y > \Delta)/\partial a = \partial a\Omega(\Delta|a)/\partial a = 1 - 2\omega a(1-\frac{\Delta}{\bar{K}})$. We show later in the proof of Proposition 2 that $1 - 2\omega a(1-\frac{\Delta}{\bar{K}}) > 0$. Finally, for part (iv), we have $d\Delta/dL = (1-p)/p\tau > 0$ and $\partial AI(a, Y > \Delta)/\partial \Delta = \partial \Omega(\Delta|a)/\partial \Delta = \omega a/\bar{K} > 0$. ***Q.E.D.***

Proof of Proposition 1

Our proof comprises three steps. We first show that without changing the auditor's reporting strategy, an increase in Y enhances audit effort;, i.e., $da^*(Y)/dY > 0$ for $Y \in [0, \Delta]$ and $Y \in (\Delta, \bar{K}]$. We then prove that a slight increase in Y that shifts the auditor's reporting strategy does not influence the auditor's effort decision; i.e., $a^*(\Delta^-) = a^*(\Delta^+)$. Finally, we demonstrate that the largest audit effort when there is no restriction on NAS in the second-best case is smaller than the audit effort in the first-best case; i.e., $a^*(\bar{K}) < a^{FB}$.

Step 1

Differentiating $a^*(Y|Y > \Delta)$ with respect to Y yields

$$\frac{da^*(Y|Y > \Delta)}{dY} = \frac{[\phi+(1-\phi)p]\omega\left(1-\frac{Y}{\bar{K}}\right)}{\frac{c}{\tau}+2(1-\phi)p\omega\Delta\left(1-\frac{\Delta}{2\bar{K}}\right)} > 0.$$

For the case of $Y \leq \Delta$, using the first-order condition for the auditor's decision problem:

$$\Pi_a(a, Y|Y \leq \Delta) = -ca + \tau\{(1-\phi)p\Delta + [\phi + (1-\phi)p(1-2a))]\omega Y\left(1-\frac{Y}{2\bar{K}}\right) = 0,$$

we obtain

$$\Pi_{aa}(a, Y|Y \leq \Delta) = -c - 2\tau(1-\phi)p\omega Y\left(1-\frac{Y}{2\bar{K}}\right) < 0,$$

$$\Pi_{aY}(a, Y|Y \leq \Delta) = \tau[\phi + (1-\phi)p(1-2a)]\omega\left(1-\frac{Y}{\bar{K}}\right) > 0,$$

where the last inequality follows from $\phi+(1-\phi)p(1-2a)>\phi-(1-\phi)p>0$. It follows that

$$\frac{da^*(Y|Y\leq\Delta)}{dY}=-\frac{\Pi_{aY}(a,Y|Y\leq\Delta)}{\Pi_{aa}(a,Y|Y\leq\Delta)}=\frac{[\phi+(1-\phi)p(1-2a^*(Y|Y\leq\Delta))]\omega\left(1-\frac{Y}{\overline{K}}\right)}{\frac{c}{\tau}+2(1-\phi)p\omega Y\left(1-\frac{Y}{2\overline{K}}\right)}>0.$$

Step 2

From the expression of $a^*(Y)$, one can easily see that $a^*(\Delta^-)=a^*(\Delta^+)$. Together with the fact that $da^*(Y)/dY>0$ for $Y\in[0,\ \Delta]$ and $Y\in(\Delta,\ \overline{K}]$, the equilibrium audit effort in the second-best case (i.e., $a^*(Y)$) is increasing in *Y*.

Step 3

We have shown that the equilibrium audit effort in the second-best case (i.e., $a^*(Y)$) is increasing in *Y*. To show that the equilibrium audit effort in the second-best is smaller than that in the first-best case, we need only to show that $a^*(\overline{K})<a^{FB}$. Using equations (18) and (19) with $Y=\overline{K}$, the result that $a^*(\overline{K})<a^{FB}$ follows from the assumptions that $\tau\leq 1$ and $I-pR>p\overline{K}>p\Delta$. ***Q.E.D.***

Proof of Corollary 1

The proof is straightforward and thus is omitted for brevity. ***Q.E.D.***

Proof of Proposition 2

The proof comprises three steps. We first show that $dAQ(a^*(Y),Y)/dY>0$ for $Y\in[0,\ \Delta]$ and $Y\in(\Delta,\ \overline{K}]$. We then demonstrate that $AQ(a^*(\Delta^-),\Delta^-)>AQ(a^*(\Delta^+),\Delta^+)$. Finally, we prove that $AQ(a^*(\Delta),\Delta)<(>)AQ(a^*(\overline{K}),\overline{K})$ is equivalent to $L<(>)\hat{L}$, where $\hat{L}\in(0,\ \overline{L})$.

Step 1

Notice that

$$\frac{dAQ(a^*(Y), Y|Y \le \Delta)}{dY} = \frac{da^*(Y|Y \le \Delta)}{dY},$$

$$\frac{dAQ(a^*(Y), Y|Y > \Delta)}{dY} = \frac{d[a^*(Y|Y > \Delta)\,\Omega\,(\Delta|a^*(Y|Y > \Delta))]}{dY}$$

$$= \left[1 - 2\,\omega\, a^*(Y|Y > \Delta)\left(1 - \frac{\Delta}{\overline{\overline{K}}}\right)\right]\frac{da^*(Y|Y > \Delta)}{dY}.$$

From the proof of Proposition 1, we have $da^*(Y)/dY > 0$ for all Y; therefore, we obtain that $dAQ(a^*(Y), Y|Y \le \Delta)/dY > 0$. In order to prove that $dAQ(a^*(Y), Y|Y > \Delta)/dY > 0$, we need to show that $1 - 2\,\omega\, a^*(Y|Y > \Delta)\left(1 - \frac{\Delta}{\overline{K}}\right) > 0$, or equivalently, $a^*(Y|Y > \Delta) < \frac{1}{2\,\omega\left(1 - \frac{\Delta}{\overline{K}}\right)}$. Since $a^*(Y|Y > \Delta) < a^*(\overline{K})$, it suffices to show that $a^*(\overline{K}) < \frac{1}{2\,\omega\left(1 - \frac{\Delta}{\overline{K}}\right)}$.

Using (19), we have

$$\frac{1}{2\,\omega\left(1 - \frac{\Delta}{\overline{\overline{K}}}\right)} - a^*(\overline{K}) \propto \frac{c}{\tau} + (1-\phi)p\omega\frac{\Delta^2}{\overline{\overline{K}}} - [\phi + (1-\phi)p]\omega^2\overline{K}\left(1 - \frac{\Delta}{\overline{\overline{K}}}\right).$$

Let $a^{SP}(\overline{K})$ be the audit effort chosen by the social planner who maximizes $V(a, \overline{K})$ in Equation (17). Note that $a^{SP}(\overline{K})$ is different from $a^{FB}(\overline{K})$ because of the presence of auditor independence concern, and is given by

$$a^{SP}(\overline{K}) = \frac{(1-\phi)(I - pR) + \frac{1}{2}[\phi + (1-\phi)p]\omega\overline{K}}{c + 2(1-\phi)(I - pR)\omega\left(1 - \frac{\Delta}{\overline{\overline{K}}}\right) + (1-\phi)p\omega\frac{\Delta^2}{\overline{K}}}.$$

Also note that, by definition, we have $a^{FB}(\overline{K}) > a^{SP}(\overline{K}) > a^*(\overline{K})$.

After going through tedious calculations, we obtain

$$a^{SP}(\overline{K}) - a^*(\overline{K}) \propto \left(\frac{c}{\tau} - c\right)\left\{\frac{1}{2}[\phi + (1-\phi)p]\,\omega\overline{K} + (1-\phi)p\Delta\right\}$$
$$+ (1-\phi)[I - p(R+\Delta)]\left\{\frac{c}{\tau} + (1-\phi)p\omega\frac{\Delta^2}{\overline{\overline{K}}} - [\phi + (1-\phi)p]\omega^2\overline{K}\left(1 - \frac{\Delta}{\overline{\overline{K}}}\right)\right\}.$$

Since $a^{SP}(\overline{K}) > a^*(\overline{K})$ and the first term above is zero when τ equals to 1, it must be that $\frac{c}{\tau} + (1-\phi)p\omega\frac{\Delta^2}{\overline{K}} - [\phi + (1-\phi)p]\omega^2\overline{K}\left(1 - \frac{\Delta}{\overline{K}}\right) > 0$; therefore, $dAQ(a^*(Y), Y|Y > \Delta)/dY > 0$.

Step 2

The difference between $AQ(a^*(\Delta^-),\Delta^-)$ and $AQ(a^*(\Delta^+),\Delta^+)$ is determined as

$$AQ(a^*(\Delta^-),\Delta^-) - AQ(a^*(\Delta^+),\Delta^+) = a^*(\Delta^-) - a^*(\Delta^+)\,\Omega\,(\Delta|a^*(\Delta^+).$$

From the proof of Proposition 1, we have $a^*(\Delta^-) = a^*(\Delta^+)$. Together with the fact that $\Omega\,(\Delta|a^*(\Delta^+)) < 1$, we obtain that $AQ(a^*(\Delta^-),\Delta^-) > AQ(a^*(\Delta^+),\Delta^+)$.

Step 3

After going through tedious calculations, we obtain

$$AQ(a^*(\bar{K}),\bar{K}) - AQ(a^*(\Delta),\Delta) = a^*(\bar{K})\Omega(\Delta\,|\,a^*(\bar{K})) - a^*(\Delta) = \frac{\omega(\bar{K}-\Delta)}{\bar{K}}(\Upsilon^2 - a^{*2}(\bar{K})),$$

where $\Upsilon \equiv \sqrt{\dfrac{\frac{1}{2}[\phi+(1-\phi)p](\bar{K}-\Delta)}{\frac{c}{\tau}+2(1-\phi)p\omega\Delta(1-\frac{\Delta}{2\bar{K}})}} \geq 0.$

Observe that $Sign\left(AQ(a^*(\bar{K}),\bar{K}) - AQ(a^*(\Delta),\Delta)\right) = Sign\left(\Upsilon - a^*(\bar{K})\right)$ except when $\Delta \to \bar{K}$, where $AQ(a^*(\bar{K}),\bar{K}) - \lim_{\Delta\to\bar{K}} AQ(a^*(\Delta),\Delta) = 0$. Also, observe that

$$\lim_{\Delta\to\bar{K}^-}(\Upsilon - a^*(\bar{K})) = -\lim_{\Delta\to\bar{K}^-} a^*(\bar{K}) < 0,$$

$$\lim_{\Delta\to 0}(\Upsilon - a^*(\bar{K})) = \sqrt{\tfrac{\tau}{2c}[\phi+(1-\phi)p]\omega\bar{K}} - \tfrac{\tau}{2c}[\phi+(1-\phi)p]\omega\bar{K}.$$

Given Condition (6) that $c > \underline{c} \equiv (1-\phi)[I - p(R+\frac{1}{2}\omega\bar{K})] + \frac{1}{2}\phi\omega\bar{K} > \frac{1}{2}[\phi+(1-\phi)p)]\omega\bar{K}$, where the last inequality follows Condition (1) that $I - pR > p\bar{K}$, we have $\frac{\tau}{2c}[\phi+(1-\phi)p]\omega\bar{K} < 1$; therefore, $\lim_{\Delta\to 0}(\Upsilon - a^*(\bar{K})) = \sqrt{\frac{\tau}{2c}[\phi+(1-\phi)p]\omega\bar{K}} - \frac{\tau}{2c}[\phi+(1-\phi)p]\omega\bar{K} > 0$. Furthermore, we have

$$\frac{\partial\Upsilon}{\partial\Delta} = -\frac{1}{2\Upsilon}\frac{[\phi+(1-\phi)p][\frac{\tau}{2c}+\frac{1}{2}(1-\phi)p\omega(\frac{\bar{K}^2+(\bar{K}-\Delta)^2}{\bar{K}})]}{[\frac{\tau}{2c}+2(1-\phi)p\omega\Delta(1-\frac{\Delta}{2\bar{K}})]^2} < 0,$$

$$\frac{da^*(\bar{K})}{d\Delta} = \frac{(1-\phi)p[1-2\omega a^*(\bar{K})(1-\frac{\Delta}{\bar{K}})]}{\frac{c}{\tau}+2(1-\phi)p\omega\Delta(1-\frac{\Delta}{2\bar{K}})} > 0,$$

where the last inequality follows from the proof in Step 1. Hence, we obtain that $\Upsilon - a^*(\bar{K})$ is a decreasing function of Δ. By the Intermediate Value Theorem, there must exist a unique point, $\hat{\Delta} \in (0, \bar{K})$, such that $\Upsilon - a^*(\bar{K}) > 0$ if, and only if, $\Delta < \hat{\Delta}$. Moreover, since $\Delta \equiv (1-p)L/p\tau$, we have $\Delta \to 0$ when $L \to 0$ and $\Delta \to \bar{K}$ when $L \to \bar{L} \equiv p\tau\bar{K}/(1-p)$. Thus, we can conclude that there must exist a unique point, $\hat{L} \in (0, \bar{L})$, such that $AQ(a^*(\bar{K}), \bar{K}) - AQ(a^*(\Delta), \Delta) > 0$ if, and only if, $L < \hat{L}$. ***Q.E.D.***

Proof of Proposition 3

We proceed our proof in three steps. We first show that $dV(a^*(Y), Y)/dY > 0$ for $Y \in [0, \Delta)$ and $Y \in (\Delta, \bar{K}]$. We then demonstrate that $V(a^*(\Delta^-), \Delta^-) > V(a^*(\Delta^+), \Delta^+)$. Finally, we prove that $L < \bar{L}$ is a sufficient condition for $V(a^*(\Delta), \Delta) < V(a^*(\bar{K}), \bar{K})$.

Step 1

The value of the audit can be determined as $V(a^*(Y), Y) = A(a^*(Y), Y) + \Pi(a^*(Y), Y)$. Using Equations (10), (12), and (14), we express the value of the audit to the firm as $A(a^*(Y), Y) = (1-\phi)AQ(a^*(Y), Y)(I - pR) + EL(a^*(Y), Y) + (1-\tau)ES(a^*(Y), Y) - F$. Since in equilibrium $\frac{\partial \Pi(a^*(Y),Y)}{\partial a^*(Y)} = 0$, we obtain that $\frac{\partial V(a^*(Y),Y)}{\partial a^*(Y)} = \frac{\partial A(a^*(Y),Y)}{\partial a^*(Y)}$. It follows that

$$\frac{dV(a^*(Y),Y)}{dY} = \frac{\partial A(a^*(Y),Y)}{\partial a^*(Y)} \frac{da^*(Y)}{dY} + \frac{\partial V(a^*(Y),Y)}{\partial Y}.$$

Using

$$\frac{\partial \Pi(a^*(Y),Y)}{\partial a^*(Y)} = -ca^*(Y) - \frac{\partial EL(a^*(Y),Y)}{\partial a^*(Y)} + \tau \frac{\partial ES(a^*(Y),Y)}{\partial a^*(Y)} = 0,$$

we rewrite $\partial ES(a^*(Y), Y)/\partial a^*(Y)$ as

$$\frac{dES(a^*(Y),Y)}{da^*(Y)} = \frac{1}{\tau} ca^*(Y) - (1-\phi)p\Delta \frac{\partial AQ(a^*(Y),Y)}{\partial a^*(Y)}.$$

Then, we can write $\partial A(a^*(Y), Y)/\partial a^*(Y)$ as

$$\frac{\partial A(a^*(Y),Y)}{\partial a^*(Y)} = (1-\phi)(I - pR - \tau p\Delta) \frac{\partial AQ(a^*(Y),Y)}{\partial a^*(Y)} + (1-\tau) \frac{\partial ES(a^*(Y),Y)}{\partial a^*(Y)}$$

$$= (1-\phi)[I - p(R + \Delta)] \frac{\partial AQ(a^*(Y),Y)}{\partial a^*(Y)} + \frac{1-\tau}{\tau} ca^*(Y).$$

As shown in the proof of Proposition 2, $\frac{\partial AQ(a^*(Y))}{\partial a^*(Y)} > 0$ for $Y \in [0, \Delta)$ and $Y \in (\Delta, \bar{K}]$, we have $\frac{\partial A(a^*(Y),Y)}{\partial a^*(Y)} > 0$ for $Y \in [0, \Delta)$ and $Y \in (\Delta, \bar{K}]$.

Next, it can be easily shown that $\frac{\partial AQ(a^*(Y),Y)}{\partial Y} = 0$ and $\frac{\partial ES(a^*(Y),Y)}{\partial Y} > 0$ for $Y \in [0, \Delta)$ and $Y \in (\Delta, \bar{K}]$, so $\frac{\partial V(a^*(Y),Y)}{\partial Y} > 0$ for $Y \in [0, \Delta)$ and $Y \in (\Delta, \bar{K}]$. Therefore, for $Y \in [0, \Delta)$ and $Y \in (\Delta, \bar{K}]$, since $\frac{\partial A(a^*(Y),Y)}{\partial a^*(Y)} > 0$, $\frac{da^*(Y)}{dY} > 0$, and $\frac{\partial V(a^*(Y),Y)}{\partial Y} > 0$, we have $\frac{dV(a^*(Y),Y)}{dY} > 0$.

Step 2

Recall that when $Y = \Delta$, the auditor is indifferent between suppressing unfavorable audit evidence and truthfully reporting it, implying that $\Pi(a^*(\Delta^-), \Delta^-) = \Pi(a^*(\Delta^+), \Delta^+)$. It then follows that $V(a^*(\Delta^-), \Delta^-) - V(a^*(\Delta^+), \Delta^+) = A(a^*(\Delta^-), \Delta^-) - A(a^*(\Delta^+), \Delta^+)$.

As shown in the proof of Proposition 1, $a^*(\Delta^+) = a^*(\Delta^-)$. Then, using Equations (17) and (19), the difference between $V(a^*(\Delta^-), \Delta^-)$ and $V(a^*(\Delta^+), \Delta^+)$ is determined as

$$V(a^*(\Delta^-), \Delta^-) - V(a^*(\Delta^+), \Delta^+) = (1-\phi)\omega[a^*(\Delta)]^2\left(1 - \frac{\Delta}{\bar{K}}\right)[I - p(R+\Delta)] > 0.$$

Step 3

Let $Z \equiv V(a^*(\bar{K}), \bar{K}) - V(a^*(\Delta), \Delta))$. After going through tedious calculations, we obtain

$$\begin{aligned} Z = {} & (1-\phi)[a^*(\bar{K})\Omega(\Delta \mid a^*(\bar{K})) - a^*(\Delta)]\{I - p[R + \tfrac{\tau}{2}\Delta + (1-\tau)E[X(\Delta) \mid a^*(\Delta)]]\} \\ & + (1-\phi)p\omega a^{*2}(\bar{K})(1 - \tfrac{\Delta}{\bar{K}})[\tfrac{\tau\Delta}{2} + (1-\tau)(\tfrac{\bar{K}+\Delta}{2} - E[X(\Delta) \mid a^*(\Delta)])] \\ & + \{\tfrac{\tau}{2}[\phi + (1-\phi)p] + (1-\tau)[\phi + (1-\phi)(1 - a^*(\bar{K}))p]\}(E[K \mid a^*(\bar{K})] - E[X(\Delta) \mid a^*(\Delta)]). \end{aligned}$$

Since $\frac{\bar{K}+\Delta}{2} > \Delta > E[X(\Delta) \mid a^*(\Delta)]$ and $E[K \mid a^*(\bar{K})] > E[X(\Delta) \mid a^*(\Delta)]$, the second and third lines of the above expression are positive. Moreover, since $\Delta > [\frac{\tau}{2} + (1-\tau)]\Delta > \frac{\tau}{2}\Delta + (1-\tau)E[X(\Delta) \mid a^*(\Delta)]$ and $I - pR > p\bar{K} > p\Delta$, the term inside the curly bracket in the first line of the above expression is positive. Hence, $Z > 0$ as long as $AQ(a^*(\bar{K}), \bar{K}) - AQ(a^*(\Delta), \Delta) = a^*(\bar{K})\Omega(\Delta \mid a^*(\bar{K})) - a^*(\Delta) > 0$. In the proof of Proposition 2, we have shown that $AQ(a^*(\bar{K}),\bar{K})-AQ(a^*(\Delta),\Delta)>0$ if $L<\hat{L}$. As such, we have $V(a^*(\bar{K}),\bar{K})>V(a^*(\Delta),\Delta))$ if $L<\hat{L}$. ***Q.E.D.***

Proof of Corollary 2

In the proof of Proposition 2, we have shown that $Sign\left(AQ(a^*(\bar{K}),\bar{K})-AQ(a^*(\Delta),\Delta)\right)=Sign\left(\Upsilon-a^*(\bar{K})\right)$ except when $\Delta\to\bar{K}$. Moreover, $\hat{L}$ is a threshold of L such that $\Phi(L,\omega)\equiv\Upsilon-a^*(\bar{K})=0$ when $L=\hat{L}$ for a given ω. We have also shown that $\Phi(L,\omega)$ is decreasing in Δ, and thus also decreasing in L. It can also be easily shown that $\Phi(L,\omega)$ is decreasing in ω. Without loss of generality, let us assume that $0<\omega_1<\omega_2<\bar{\omega}$ and, accordingly, $\hat{L}_1$ and $\hat{L}_2$ are the respective thresholds of L such that $\Phi(L,\omega)=0$ for any given ω. We prove that $\hat{L}_1>\hat{L}_2$ by contradiction. Suppose that $\hat{L}_1<\hat{L}_2$. Since $\Phi(L,\omega)$ is decreasing in ω, we have $\Phi(\hat{L}_1,\omega_2)<0$. Moreover, since $\Phi(L,\omega)$ is decreasing in L, we have $\Phi\left(\hat{L}_2,\omega_2\right)\le\Phi\left(\hat{L}_1,\omega_2\right)<0$, which contradicts $\Phi(\hat{L}_2,\omega_2)=0$. Hence, it must be $\hat{L}_1>\hat{L}_2$. ***Q.E.D.***

REFERENCES

Antle, R., and J.S. Demski, 1991, Contracting Frictions, Regulation and the Structure of CPA Firms, *Journal of Accounting Research* 29 (Supplement): 1–24.

———, E.A. Gordon, G. Narayanamoorthy, and L. Zhou, 2006, The Joint Determination of Audit Fees, Non-Audit Fees, and Abnormal Accruals, *Review of Quantitative Finance and Accounting* 27 (3): 235–266.

Arens, A.A., and J.K. Loebbecke, 1981, *Applications of Statistical Sampling to Auditing*, Englewood Cliffs, NJ: Prentice Hall.

Arruñada, B., 1999, *The Economics of Audit Quality: Private Incentives and the Regulation of Audit and Non-audit Services*, Kluwer, Dordrecht.

Ashbaugh, H., R. LaFond, and B.W. Mayhew, 2003, Do Nonaudit Services Compromise Auditor Independence? Further Evidence, *The Accounting Review* 78 (July): 611–639.

Beck, P. J., T. J. Frecka, and I. Solomon, 1988, A Model of the Market for MAS and Audit Services: Knowledge Spillovers and Auditor-Auditee Bonding, *Journal of Accounting Literature* 7: 50–64.

Causholli, M., D.J. Chambers, and J.L. Payne, 2014, Future Nonaudit Service Fees and Audit Quality, *Contemporary Accounting Research* 31(3): 681–712.

Chan, D.K., 2016, Financial Reporting and Auditing under Alternative Damage Apportionment Rules Revisited, Working Paper, The University of Hong Kong.

———, 2018, The Effects of Internal Control Reporting Regulation on Control Quality, Compensation and Audit Effort, Working Paper, The University of Hong Kong.

———, and N. Liu, 2023, Critical Audit Matter Disclosure, Audit Quality and Investment Efficiency, *The Accounting Review*, 98(2): 97–121.

———, and S. Pae, 1998, An Analysis of the Economic Consequences of the Proportionate Liability Rule, *Contemporary Accounting Research* 15 (Winter): 457–480.

———, and K.P. Wong, 2002, Scope of Auditors' Liability, Audit Quality, and Capital Investment, *Review of Accounting Studies* 7: 97–122.

Chan, L., T.Y. Chen, S. Janakiraman, and S. Radhakrishnan, 2012, Reexamining the Relation between Audit and Non-Audit Fees: Dealing with Weak Instruments in Two-Stage Least Squares Estimation, *Journal of Accounting, Auditing and Finance* 27(3): 299–324.

Chen, Q., X. Jiang, and Y. Zhang, 2019, The Effects of Audit Quality Disclosure on Audit Effort and Investment Efficiency, *The Accounting Review* 94(4): 189–214.

Chung, H., and S. Kallapur, 2002, Client Importance, Non-Audit Fees, and Abnormal Accruals, *The Accounting Review* 78: 931–55.

Davis, L.R., D.N. Ricchiute and G. Trompeter,1993, Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients, *The Accounting Review* 68 (1): 135–150.

DeAngelo, L.E., 1981, Auditor Size and Audit Quality, *Journal of Accounting and Economics* 3 (3): 183–199.

DeFond, M.L., K. Raghunandan, and K.R. Subramanyam, 2002, Do Non-audit Service fees Impair Auditor Independence? Evidence from Going Concern Audit opinions, *Journal of Accounting Research* 40 (4): 1247–1274.

Deng, M., N.D. Melumad, and T. Shibano, 2012, Auditors' Liability, Investments, and Capital Markets: A Potential Unintended Consequence of the Sarbanes-Oxley Act, *Journal of Accounting Research* 50(5): 1179–1215.

Dopuch, N., R. King, and R. Schwartz, 2004, Contingent Rents and Auditors' Independence, *Asia-Pacific Journal of Accounting and Economics* 11: 47–70.

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

———, 1995, Incorporation and the Audit Market, *Journal of Accounting and Economics* 19: 75–114.

Ewert, R., 2004. Audit Regulation, Audit Quality, and Audit Research in a Post-Enron Era: An Analysis of Nonaudit services. In: Leutz, C., Pfaff, D. and Hopkins, A. (Eds): *The Economics and Politics of Auditing*, Oxford: 239–263.

Financial Reporting Council. 2019. Revised Ethical Standard 2019. London, UK.

Frankel, R. M., M. F. Johnson, and K. K. Nelson, 2002, The Relation between Auditors' Fees for Nonaudit Services and Earnings Management, *The Accounting Review* (October): 71–105.

Friedman, H.L., and L. Mahieux, 2021, How Is the Audit Market Affected by Characteristics of the Nonaudit Services Market? Journal of Accounting Research 59(3): 959–1020.

Gao, P., and G. Zhang, 2019, Auditing Standards, Professional Judgement, and Audit Quality, *The Accounting Review* (November): 201–225.

Hillegeist, S.A., 1999, Financial Reporting and Auditing Under Alternative Damage Apportionment Rules, *The Accounting Review* 74 (July): 347–369.

Kinney, W. R., Z. Palmrose, and S. Scholz, 2004, Auditor Independence, Nonaudit Services and Restatements: Was the U.S. Government Right? *Journal of Accounting Research* (June): 561–588.

Kornish, L.J., and C.B. Levine, 2004, Discipline with Common Agency: The Case of Audit and Nonaudit Services, *The Accounting Review* 79 (January): 173–200.

Laux, V., and P. Newman, 2010, Auditor Liability and Client Acceptance Decisions, *The Accounting Review* 85 (2): 261–285.

Liu, X., and D. Simunic, 2005, Profit Sharing in an Auditing Oligopoly, *The Accounting Review* 80 (2): 677–702.

Lowe, D.J., M.A. Geiger, and K. Pany, 1999, The Effects of Internal Audit Outsourcing on Perceived External Auditor Independence, *Auditing: A Journal of Practice and Theory* 18 (Supplement): 7–26.

Lu, T., and H. Sapra, 2009, Auditor Conservatism and Investment Efficiency, *The Accounting Review* 84 (6): 1933–1958.

Mahieux, L., 2022, Auditors' Incentives and Audit Quality: Non-Audit Services versus Contingent Audit Fees, *European Accounting Review*: 1–37.

Pae, S., and S. Yoo, 2001, Strategic Interaction in Auditing: An Analysis of Auditors' Legal Liability, Internal Control System Quality, and Audit Effort, *The Accounting Review* 76: 333–356.

Patterson, E., and R. Smith. 2007, The Effects of Sarbanes-Oxley on Auditing and Internal Control Strength. *The Accounting Review* 82: 427–455.

Raghunandan, K., W. Read, and S. Whisenant, 2003, Initial Evidence on the Association Between Non-Audit Fees and Restated Financial Statements, *Accounting Horizons* 17: 223–34.

Robinson, D., 2008, Auditor Independence and Auditor-Provided Tax Service: Evidence from Going-Concern Audit Opinions Prior to Bankruptcy Filings, *Auditing: A Journal of Practice & Theory* 27: 31–54.

Ruddock, C.M.S., S.J. Taylor, and S.L. Taylor, 2006, Non-Audit Services and Earnings Conservatism: Is Auditor Independence Impaired? *Contemporary Accounting Research* 23: 701–746.

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

Swanger, S.L., and E.G. Chewning, Jr., 2001, The Effect of Internal Audit Outsourcing on Financial Analysts' Perceptions of External Auditor Independence, *Auditing: A Journal of Practice and Theory* 20 (September): 115–129.

Wu, M.G.H., 2006, An Economic Analysis of Audit and Non-Audit Services: The Tradeoff between Competition Crossovers and Knowledge Spillovers, *Contemporary Accounting Research* 23 (2): 527–554.

Ye, M., and D. Simunic, 2022, The Impact of PCAOB Type of Regulations on Auditors under Different Legal Systems, *Journal of Accounting, Auditing and Finance*, forthcoming.

Figure 1: Time Line

0 — The firm claims its type is good and hires an auditor to attest to its claim by paying a market-determined flat audit fee, F.

1 — The auditor exerts effort to collect audit evidence on the firm's type, privately observes the value of a consulting opportunity, K, that might arrive at date 3, and then decides whether to issue an unmodified opinion, $\hat{G}$, or a modified opinion, $\hat{B}$.

Given the audit report, the investors invest I in the firm's project if, and only if, the firm receives the unmodified report, $\hat{G}$. If investors do not invest, the game ends with the firm being liquidated.

2 — The firm realizes cash flows from its project. If the project succeeds, it yields R. If the project fails, it yields zero and the firm's true type is revealed.

The investors sue the auditor if, and only if, the firm is revealed to be a bad type. The court assigns damage awards to be paid by the auditor.

3 — An expansion option is available to the firm if the project succeeded.

The firm hires the incumbent auditor to provide consulting services for the expansion option.

The expansion option yields K, and the firm pays the auditor τ proportion of the option yield as consulting fee.

Figure 2: Game Tree When $Y > \Delta$

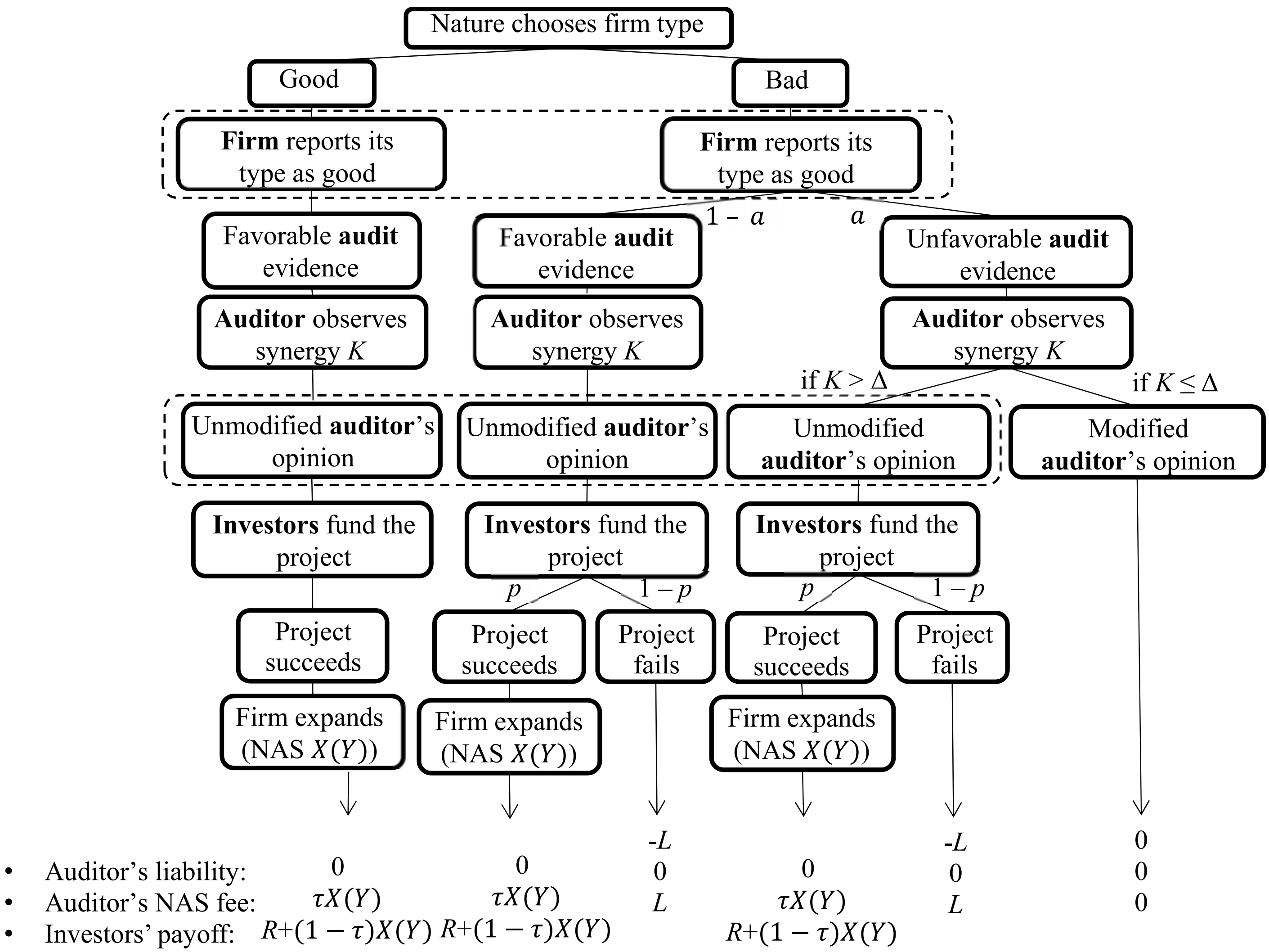
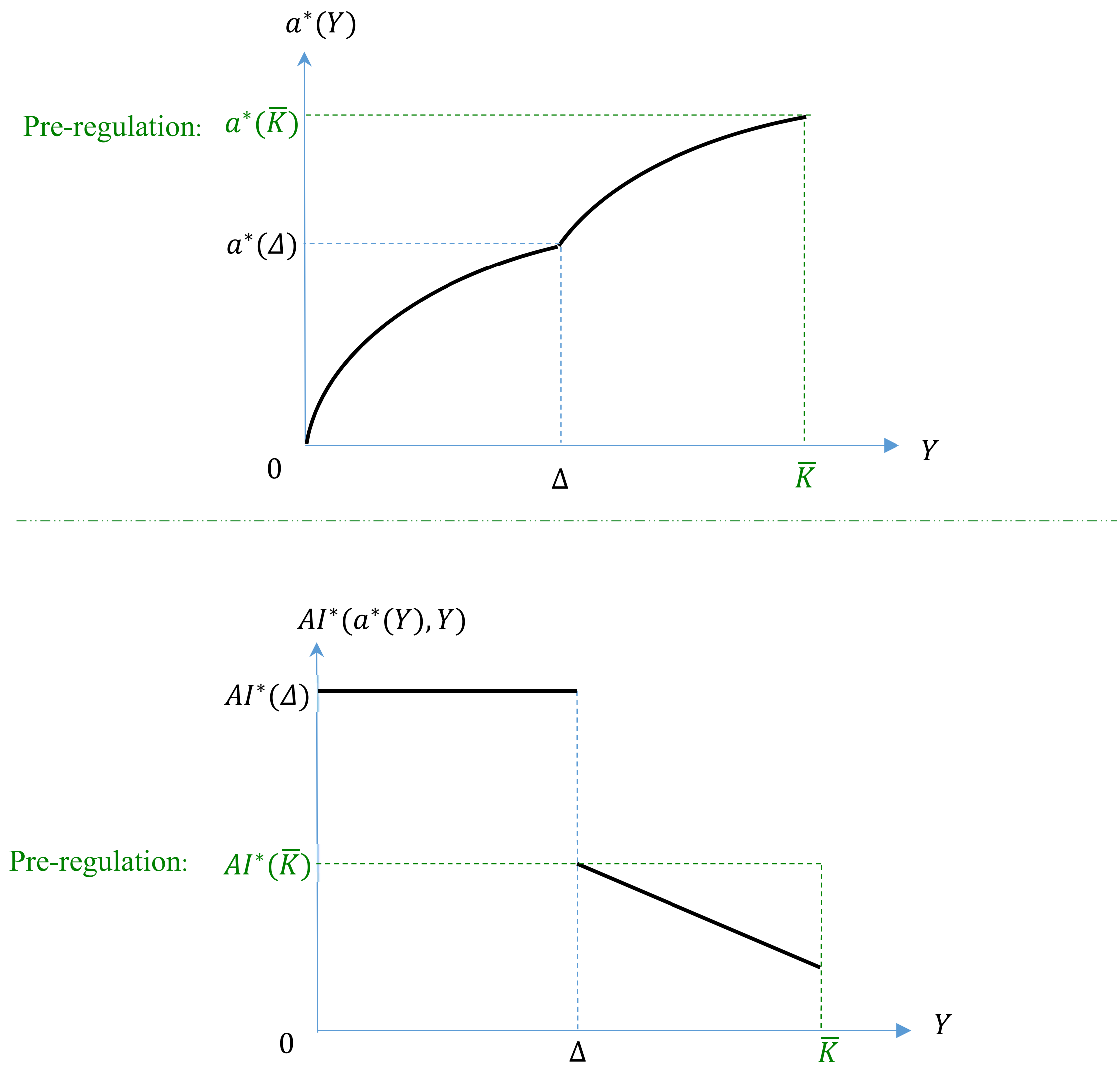


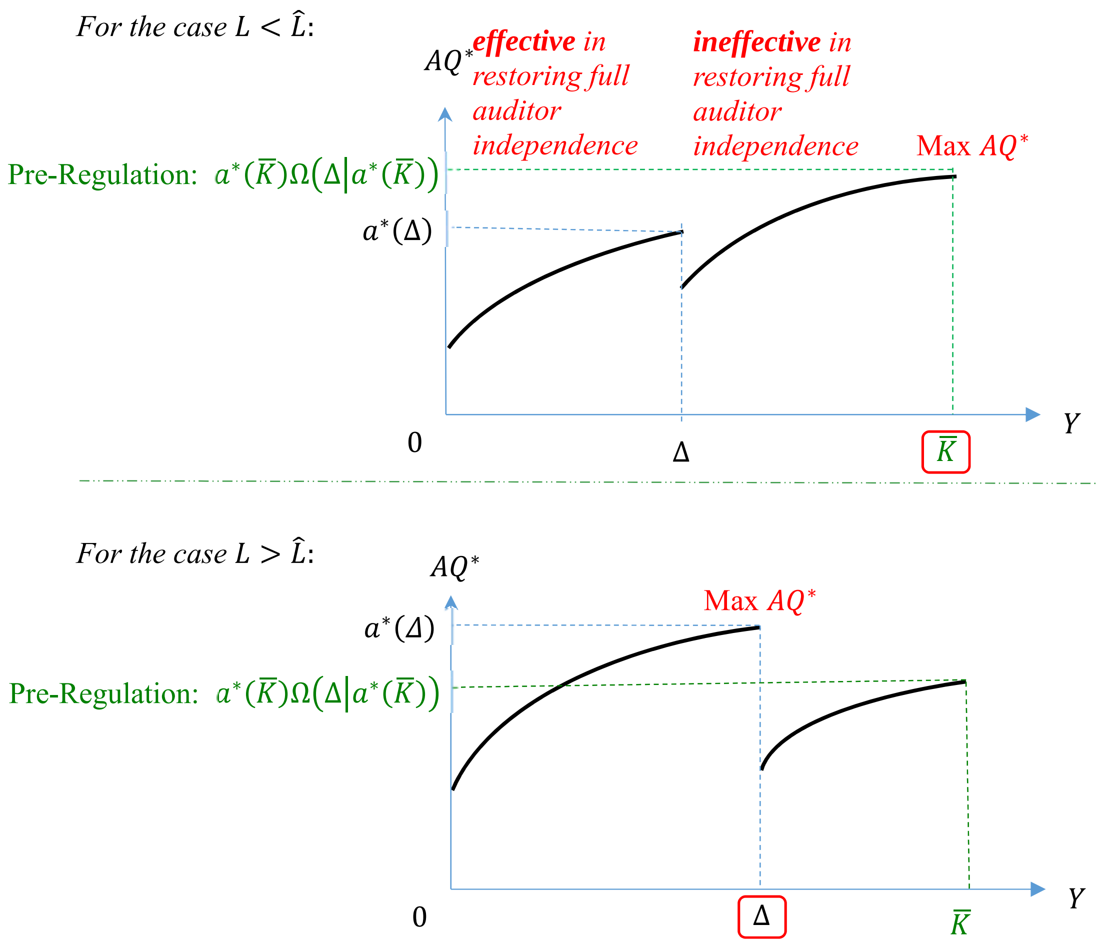
Figure 3: The Impact of a Restriction of NAS on Audit Effort and Auditor Independence



The equilibrium audit effort $a^*(Y)$ is increasing in Y. It increases faster at $Y = \Delta^+$ than at $Y = \Delta^-$.

The equilibrium auditor independence $AI^*(a^*(Y), Y)$ is independent of Y when $Y \leq \Delta$ and decreasing in Y when $Y > \Delta$.

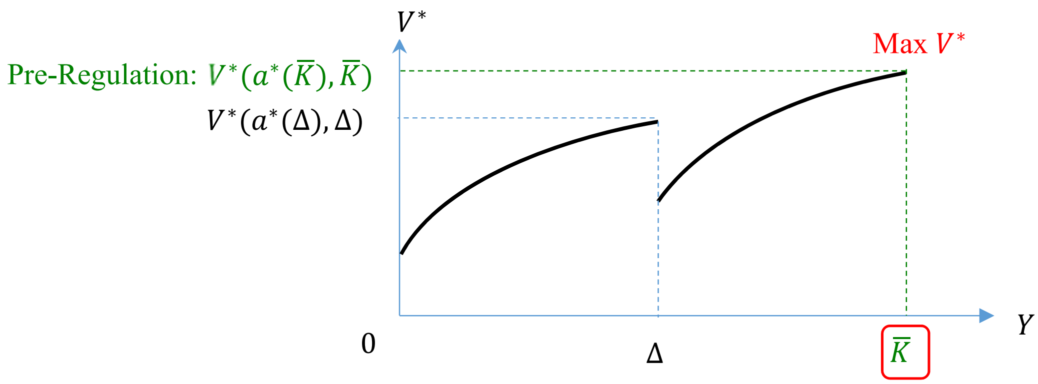
Figure 4: The Impact of a Restriction of NAS on Audit Quality



Audit quality $AQ(a^*(Y), Y)$ increases for $Y \in [0, \Delta)$, jumps down at $Y = \Delta$, and then increases for $Y \in (\Delta, \bar{K}]$. There exists a unique threshold, $\hat{L} \in (0, \bar{L})$, such that $AQ(a^*(\Delta), \Delta) < AQ(a^*(\bar{K}), \bar{K})$ if, and only if, L is below $\hat{L}$.

Figure 5: The Impact of a Restriction of NAS on the Value of the Audit

For the case $L < \hat{L}$:



The value of the audit $V^*(a^*(Y), Y)$ increases for $Y \in [0, \Delta)$, jumps down at $Y = \Delta$, and then increases for $Y \in (\Delta, \overline{K}]$. A sufficient condition for $V^*(a^*(\Delta), \Delta) < V^*(a^*(\overline{K}), \overline{K})$ is $L < \hat{L}$.

Table 1: Notation

I	Required capital investment to undertake the firm's project.
R	Return of the project in the case of project success.
$\{G,\ B\}$	Type of firm, where G represents a good type and B represents a bad type.
ϕ	Common prior probability that the firm type is G.
p	The probability that type B firm's project succeeds.
F	Competitively determined audit fee.
$a \in [0,\ 1]$	The auditor's effort.
c	Cost parameter of audit effort.
$\{g,\ b\}$	Type of audit evidence privately observed by the auditor, where g represents a good type and b represents a bad type.
$\{\hat{G},\ \hat{B}\}$	Set of audit reports on the firm type, where $\hat{G}$ represents that the auditor agrees with the firm's report and issues an unmodified opinion, and $\hat{B}$ represents that the auditor disagrees with the firm's report and issues a modified opinion.
$K \in [\underline{K},\ \overline{K}]$	Benefit from the joint provision of audit and NAS to the same client for implementing the expansion option after the firm's project has been undertaken.
$\Omega(K \mid a) = 1 - \omega a(1 - \frac{K}{\overline{K}})$	The cumulative distribution function of K when the audit effort is a.
ω	Measure of the effectiveness of audit effort in generating a higher K.
$X(Y)$	The actual amount of the legally allowed benefit that can be derived from the joint provision of audit and NAS, $X(Y) = \min\{K,\ Y\}$, where $Y \le \overline{K}$ is the maximum amount of the benefit that is implied by the NAS restriction.
$\tau \in (0,\ 1)$	The fraction of $X(Y)$ that the auditor receives as his/her consulting fee.
L	The auditor's liability payment to the investors.
$\Delta \equiv (1 - p)L / p\tau$	The threshold of K such that the auditor will suppress the unfavorable audit evidence b if, and only if, $K > \Delta$.
AI	Auditor independence.
AQ	Audit quality.
ER	The investors' net expected return from the project without the expansion option.
ES	The investors' expected synergy from the joint provision of audit and NAS if the expansion option is exercised.
EL	The investors' expected payoff from litigation against the auditor.
U	The investors' expected continuance value of the firm given audit report $\hat{G}$.
A	The ex ante value of the audit to the firm.
Π	The auditor's ex ante expected payoff.
$V = A + \Pi$	The ex ante value of the audit.