Review of Accounting Studies, 10, 59–92, 2005

Auditor Size, Market Segmentation and Litigation Patterns: A Theoretical Analysis

SASSON BAR-YOSEF mssasson@huji.ic.il
School of Business, The Hebrew University, Jerusalem, Israel

BHARAT SARATH* bharat_sarath@baruch.cuny.edu
Zicklin School of Business, Baruch College, New York, NY

Abstract. We provide a theoretical rationale for the observed audit industry structure where well-capitalized auditors hold an extremely large market share. Our analysis focuses on the economics of trading in an adverse selection market where audit quality is unobservable. We show that concentration of market share can arise even if well-capitalized auditors have no relative advantage with regard to supplying high-quality audits, and that the strategy of attracting a narrow base of high-margin clients is typically unsustainable in rational expectations equilibrium. Other results derived from our analysis of strategic competition for clients also conform (qualitatively) with empirical findings regarding audit fee structures and litigation rates. In particular, we show that better-capitalized auditors get a dominant market share, produce more accurate reports and are more profitable. In addition, we show that the imposition of high minimum standards increases the market power of wealthy auditors, even though smaller auditors can potentially provide the same level of audit quality at lower fees. All these results are demonstrated within a framework that endogenizes both a securities trading market and profit-maximizing auditors who strategically compete for clients.

Keywords: auditor-choice, audit-quality, litigation, market-shares

JEL Classification: C72, D43, D82, K23, K41, L15

Well-capitalized audit firms (such as the current "Big Four") have traditionally held very large market-shares in the audit services sector.[1] The rationale for such concentration has not been fully explored in the theoretical literature. The purpose of this study is to show how basic information asymmetries in securities markets combined with the institutional features underlying the Securities Act of 1934 lead to an outcome where market share is concentrated in the hands of wealthy audit partnerships. Two economic factors are endogenized in our analysis: (i) competition for clients by multiple auditors and (ii) a market where the securities of these client-firms are traded. The integration of these two features distinguishes our study from earlier papers that have concentrated on the behavior of a single auditor in the context of trading and securities litigation (see for instance, Dye, 1993; Menon and Williams, 1994).

We consider two auditors who are equally efficient in providing accurate public reports, but, as a consequence of random events, differ in wealth levels, and thus, in their capacity to pay damages under common law provisions and under statutory provisions of the 1933 and 1934 Securities Act (see Ratner, 1988 for details of the

*Corresponding author

Securities Acts).[2] The auditors strategically compete for client-firms who are differentiated by the level of business risk.[3] Each auditor tries to maximize own profits by attracting clients through appropriate choices of stated fees and audit quality. We demonstrate that the best competitive strategy for less well-capitalized auditors is to concentrate on higher-risk firms. In addition, we show how increasingly complex GAAP standards may benefit wealthy auditors in attaining greater (relative) levels of profitability, and, by implication, to maintain (and perhaps increase) their wealth over a long horizon.[4] We emphasize that our analysis is conducted within a framework where auditors are initially endowed with differential levels of wealth; our results demonstrate how these differences may be sustained or widened through strategic competition.

The equilibrium market-share allocations derived through our analysis show that the wealthy auditor typically captures a large share of the market. In contrast, an alternative structure (found in many service industries) consists of small market-share high-quality firms that we characterize as "boutiques."[5] We argue that boutique auditors who attract only high-type clients are unlikely to survive in equilibrium. In order to survive, such boutique auditors have to convince users (clients and investors) that their audit quality is very high and that their client-firms are of the best type. However, with a portfolio of high-quality firms, the auditor faces very little expected litigation. Because audit quality is both costly to supply and difficult to verify, a reduced threat of litigation makes it optimal for the boutique auditor to bank on their reputation rather than supply high-quality audits. Investors anticipate potential quality reductions and discount the boutique auditor's reports. However, if the boutique auditor's reports are discounted, high-type firms have no incentives to stay with this auditor and pay high fees. Consequently, the strategy of high margins and small market share is unsustainable in equilibrium.[6]

The provision of audit services under moral hazard has been analyzed in many papers (some examples are Antle, 1982; Baiman et al., 1987; Antle and Demski, 1991) and the effect of auditor wealth on reporting behavior (for a single auditor) is studied in Dye (1993). However, in these earlier studies, *ex post* penalties imposed on auditors are treated as contracting mechanisms. In contrast, in this paper, the level of penalties on auditors is determined as an endogenous outcome of competition for clients. Competition for clients through audit quality choices is studied in Ronen (1994), and pricing strategies for attracting clients have been discussed in several studies (see, for example, De Angelo, 1981). Another strand of literature investigates the effects of auditor choice on resulting market prices (see Beatty, 1989 for equity and Mansi et al., 2003 for the effect on bond prices). We also incorporate this feature into our analysis and allow market prices to be affected by auditor choice based on *investor perceptions* regarding audit quality and the types of client-firms choosing a particular auditor. All these features of price and quality competition are synthesized in this paper into a single economic framework, where auditors function under moral hazard and are subject to liability for errors.

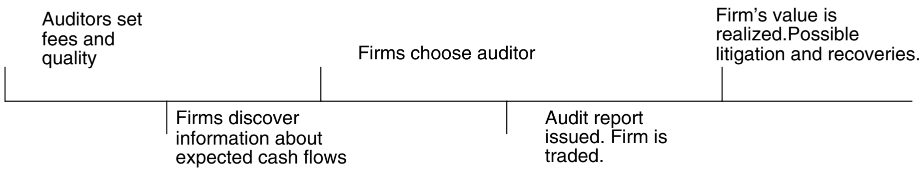
Securities trading (of equities) involves a very basic conflict between incumbent shareholders (in our context, "client-firms") who would like to sell "high" and potential investors who would prefer to buy "low." This conflict is further exacerbated by the fact that incumbent shareholders may be better informed about the security that they hold (and wish to sell). Intra-investor conflicts combined with the unobservability of the accuracy of audited reports leads to complex interactions involving simultaneous considerations of three sets of strategic issues: (1) client-firms who possess private information about their own value and seek to sell their securities at the highest possible price such as in an Initial Public Offering;[7] (2) the auditor who is hired by the firm and audits, for a predetermined fee, the financial reports while maximizing own profit (i.e., by maximizing market share, given audit fees, audit costs and potential liability); (3) investors who price the firm competitively (including potential recoveries from the auditor) based on the report received from the auditor and their perceptions regarding the accuracy of the audit report. We model these interactions within a market setting where two auditors compete for client-firms through strategic fee and quality choices.

Demand from client-firms for auditor services is driven by a need to sell their securities at the highest possible price (net of audit fees). The expected price of a security depends both on the audit report and market perceptions of the quality of the auditor. Therefore, auditors try to maximize profits through the strategic choice of audit fees and quality, taking into account the market's beliefs about their audit quality, which, in the context of our model, may be termed as their "reputation." The two strategic choices made by auditors are: (i) the resources devoted to generating the report (and thus establishing the "quality" of the report); (ii) fees that will compensate them for both the cost of producing the report and possible losses due to litigation, but will nevertheless be attractive to as many clients as possible. Our analysis emphasizes the client-demand side, whereas production aspects of audits are simplified through an assumption that all audit firms face identical costs for generating quality audits.[8]

Our main analysis concentrates on the case where two auditors with differential capacities to pay damages compete for client firms. The key goals are to show why boutique auditors are untenable in equilibrium and to explain why wealthier auditors (endogenously) capture large market shares (Propositions 5 and 6). In establishing these results, we begin by demonstrating that the market segments along levels of client risk (Proposition 1). The next result (Proposition 2) shows that if the differences in wealth levels are large, equilibrium exists and the wealthy auditor supplies audits of greater accuracy in equilibrium. Proposition 3 shows that the auditor who supplies more accurate audits (irrespective of wealth levels) attracts low-risk client firms. Proposition 4 shows that the relationship between audit accuracy and litigation damages may be confounded by the fact that client-firms are not homogenous and that higher (per-client) damages may be paid by the auditor who supplies more accurate reports. Results on market shares are presented in Propositions 5 and 6. Proposition 5 shows that under an intuitive restriction on the cost structure, boutique auditors are not tenable and that the wealthy auditor typically captures a larger market share. Proposition 6 shows, independently of Proposition 5, that the imposition of "high" minimum audit standards confers a competitive advantage on the wealthy auditor allowing him to capture a larger share of the market.

1. The Economic Environment

The principal economic features that motivate our analysis are the interaction between litigation recoveries from auditors, competition for clients by auditors, and asymmetric information across firms and investors within the context of efficient markets. Owners of client-firms are privately informed about their prospects before they approach auditors. Audit quality cannot be directly observed, and thus, cannot be used as a basis for contracts between auditors and their clients. However, client-firms choose auditors based on their beliefs regarding the quality of the report issued by the auditors. Auditors anticipate client reactions when strategically setting fees and report-quality at levels that attract clients and maximize total profits. Market prices are set by investors who observe both the auditor employed and the report issued by that auditor. The time line of events and decisions is given below:



We develop a single period, two-auditor model of the sequence of events described in the time line above. Before presenting the structure of this model, we outline the role of auditor wealth and its effects on the strategic choices made by auditors.

1.1. Auditor Wealth and Liability Rules

In order to capture the essence of strategic competition through fee and quality choices, it suffices to consider two distinct auditors, denoted by B ("big" pockets) and S ("shallow" pockets), that compete to provide audit services to a specified set of firms. We emphasize that at this stage, B and S refer solely to the levels of wealth and not to market shares. The level of audit inputs, e.g. investment in resources and personnel that are made by the audit firm is denoted by e, and is assumed to apply uniformly across all clients.[9] The higher the choice of e, the more (statistically) accurate is the auditor's report regarding a client-firm's (uncertain) future cash flows. The cost of supplying an audit of accuracy level e is denoted by $C(e)$ and is assumed to be identical across auditors;[10] in contrast, the damages paid to investors who incur losses through reliance on erroneous audit reports differ across auditors.

The litigation exposure of each auditor type, modeled in our paper, is broadly consistent both with damage provisions of securities acts and with tort liability. In either case, the two prerequisites for a successful recovery are (i) that auditors signed an erroneous report about a security to be issued and (ii) that investors who traded this security consequently suffered a market loss. Denote the trading price of a firm (consequent to an audit report) by p and the (eventually) realized value by v. For a successful recovery, it must be the case that $(p - v) > 0$ and the expected recoveries from an auditor of type J are then denoted by $\mathscr{L}_J(e, p - v)$ where $\mathscr{L}_J$ is decreasing in the level of inputs e and increasing in the market loss $(p - v)$. "Big Pockets" (or equivalently, "Deep Pockets") ensure that for any level of audit quality and market loss, *greater recoveries are possible from the auditor with "big" pockets,* that is, for a given market loss $p - v$, $\mathscr{L}_B(e, p - v) > \mathscr{L}_S(e, p - v)$ for every level of inputs, e.[11]

1.2. Economic Interactions

We describe the sequence of events, decisions, and the information available prior to, and at each stage, of the time line and simultaneously introduce the notation used to represent these interactions.

0. Auditors are endowed with differential levels of wealth and investors form beliefs regarding the choices of inputs, ε^i_j, fees, $F^i_{j,}$ and client base, Σ^i_J, where J denotes the auditor's type.

1. Auditors set strategies consisting of audit inputs and audit fees, $\{e_J, F_J\}$. The chosen level of audit inputs, e_J, is observed neither by the client-firms nor by the market. Audit fees are observable by client-firms before trade takes place but not by the market until after the trade takes place. The fees are set based on rational expectations regarding the client-base that will be attracted, based on these observed fees (see stage 3 below).

2. The future cash flows of a firm of type ω are specified by a distribution $f(c|\omega)$. Firms get imperfect information, ϕ, regarding their type, ω, and hence of the distribution of their future cash flows (and hence, their intrinsic value).

3. Firms choose an auditor, J, based on: (i) the quoted audit fee, F_J; (ii) the client-firm's beliefs ε^f_j regarding the input the auditor has chosen to supply (which determines the accuracy of the audit report); (iii) their private information, ϕ. The endogenous selection of an auditor leads to a division of the market between auditors (which is consistent with the anticipation at stage 1 above). Client-firms gain by being publicly traded, but if the costs of audits are higher than these gains, these firms will prefer to go private and avoid the audit cost (only firms with publicly traded securities must hire auditors under the Securities Acts of 1933 and 1934); for this reason, we assume an upper bound on audit fees. Thus, in equilibrium, some subset of firms, Σ_B, choose Auditor B, while the complement, Σ_S, chooses to go to Auditor S.

4. The chosen auditor, J, signs-off on an audited financial report, θ. The report is generated through the following process. Auditors begin with the knowledge that

their client-firm is chosen from their (prior) client base, Σ^i_J, so their expected probability distribution of client-firm type is $f(\omega|\Sigma_J{}^i)$. Based on their actual inputs e_J, auditors update these priors. The higher the choice of e_J, the more likely it is that the updated beliefs will be close to the true type of the client firm. The audit report θ is an accurate reflection of the updated beliefs of the auditor where the degree of updating depends of the choice of e_J. The report is assumed to be non-strategic, that is, auditors report their true assessment of the client firm. Given this report, investors set the price of the firm at $p(\theta|\varepsilon^i_J, \Sigma^i_J)$, where ε^i_J denotes the investor's beliefs regarding the quality supplied by auditor J, and $\Sigma_J{}^i$ represents the prior beliefs regarding the set of firm-types that select auditor J.[12] Potential recoveries through the litigation process are part of the determination of $p(\theta|\varepsilon^i_J, \Sigma^i_J)$. Note that at this stage, client-firms who have *observed* the actual audit fees, F_J, may revise their beliefs regarding the average firm-types with auditor J. In particular, the *posterior* beliefs about the average firm-type with auditor J, denoted by Π^*_J, may differ from the prior average, Π_J (that continue to be held by investors who do not observe the audit fees).

5. The value of the firm becomes known to investors who may then initiate a lawsuit if they have suffered a loss. A key feature of our model is that the wealthy auditor faces a greater expected cost per dollar of damage.

All beliefs and decisions are sequentially rational and anticipate the optimal strategies that will be chosen by other participants (the equilibrium is Bayesian–Nash).

The overall structure of our model can be described as two linked sub-games: (1) a Nash game between auditors who compete for clients and where each auditor attempts to maximize own profit (as discussed in detail in IIIC), and (2) a Stackelberg game between auditors and client firms. The sub-game between auditors and client-firms has the auditor acting as a Stackelberg leader, who has already chosen fees and audit inputs. After *observing* the audit fee structure, based on their *beliefs* regarding audit accuracy, client-firms respond by selecting the auditor who will maximize their net payoff (i.e., the expected trading price that will be offered by investors less audit fees). Preceding this Stackelberg sub-game is the Nash game across auditors where fees and inputs are chosen so as to maximize auditor profits taking into account the entire second stage Stackelberg sub-game. At this first stage, auditors know that their choice of fees will be revealed to clients and will directly affect their market share (and profits), whereas their choice of inputs will be unobservable and will not directly affect their market share (though it will affect the level of damages, and thereby, profits). Auditors can, therefore, deviate from the perceived level of accuracy without being detected (at the time of trade); however, in equilibrium, beliefs regarding the auditor-chosen level of accuracy and the actual choices must coincide.

Formally, an equilibrium (in pure strategies) for our model consists of a specification of optimal strategies and beliefs:

Auditors'choices:	$\{\{e_B, F_B\}; \{e_S, F_S\}\}$
Firm's choice of auditor:	$\{B \text{ or } S\}$
Market prices:	$\{P_B, (\theta\|\varepsilon^i_B\Pi^i_B), P_S(\theta\|\varepsilon^i_S\Pi^i_S)\}$
Firm's beliefs:	$\{\varepsilon^f_B\Pi^f_B, \varepsilon^f_S\Pi^f_S\}$
Investor's beliefs:	$\left\{\varepsilon^i_B,\ \Pi^i_B,\ \varepsilon^i_S,\ \Pi^i_S\right\}$

where for each auditor J, ε_J denotes beliefs held by firms and investors regarding audit accuracy and Π_J the beliefs regarding the average client-firm type (i.e., $\Pi_J = E[\omega|\Sigma_J]$).[13] The beliefs regarding the set of firms choosing each auditor, Σ_B, Σ_S, may be replaced by the expected type across these sets, Π^i_B and Π^i_S, because (under risk-neutrality) only the expectations are needed to determine equilibrium prices.

The beliefs of investors and firms regarding audit inputs, $\varepsilon^f_J, \varepsilon^i_J$, may be set equal for the following reason. Rational expectations equilibria require that both these sets of beliefs coincide with the actual input choice at the end of the game. In addition, neither investors nor firms obtain any information during the game that causes them to revise their beliefs. Therefore, any initial configuration where $\varepsilon^f_J \neq \varepsilon^i_J$ cannot lead to an equilibrium so we can set $\{\varepsilon^f_B, \varepsilon^f_S\} = \{\varepsilon^i_B, \varepsilon^i_S\}$ and drop the superscripts on the beliefs, ε. Note that the information structure (developed in detail later) precludes updating of beliefs regarding auditor effort (see also footnote 23).We also write Π_J for Π^i_J because investor beliefs, by themselves, determine prices whereas client-firm beliefs, Π^f_J, play only an indirect role in the analysis. Finally, we assume that investor beliefs are common knowledge (i.e., investors beliefs are known to both firms and auditors).

We work backwards by solving first for the equilibrium market prices, $p_J(\theta|\varepsilon_J, \Pi_J)$, contingent on an arbitrary set of beliefs $\{\varepsilon_J, \Pi_J\}$. Based on these prices , beliefs, and observed fees, F_J, we can determine the optimal auditor choice made by client firms. The next step is to solve for the profit maximizing strategies of auditors given the client-firms decision rules and the last step is to ensure that the initial beliefs conform with the actual strategy choice and realized market share of each auditor. A complete illustration of this process is provided in Appendix C.

The next stage of our development is to impose more structure on the relationship between uncertain future cash flows and the audit report so as to allow a determination of the report-contingent price offered by investors.

2. The Complete Model And Equilibrium Solution

We choose a family of distributions for the future values of the firm that is sufficiently simple to allow us to solve explicitly for current prices and yet flexible enough to permit consideration of audit liability issues. Each client-firm invests in a single project that pays off one of two values in the future: x (successful) or 0 (failure) where x is some fixed positive quantity. The true probability of attaining the high value x (success) is denoted by ω and is assumed to differ across projects and is unknown at the time of investment. Each value of ω yields a firm with a different pattern of (probabilistic) future cash flows; ω is assumed to be distributed uniformly on (0,1).[14] Investing in the project endows the firm with an imperfect estimate of ω, denoted by ϕ.[15]

Investors use audit reports to form conjectures about the likely return from purchasing securities in the firm. Because all the characteristics of future value distribution are captured by a single parameter $\omega \in (0,1)$, each audit report becomes associated with a value $\theta \in (0,1)$. An audit report, θ, represents a set of financial statements which, if completely accurate, would lead investors to conclude that the firm's value in the future will be x with probability θ and 0 with probability $(1 - \theta)$.

Next, we model the audit input level e as the probability that the audit is accurate, that is, the complementary probability $(1-e)$ represents the likelihood of an audit error. Whenever the audit has no errors (with probability e), the underlying type ω is correctly identified, that is, the report $\theta = \omega$ with probability e. However with probability $(1-e)$, the audit process contains errors and the firm's type is incorrectly identified. To simplify the analysis, we assume that when there are errors in the audit process, the identification of the firm's type is "pure noise," that is, the identified type, and consequently the report, are uniformly distributed over (0,1) and independent of the true type, ω. Thus, when there is an audit failure (as happens with probability $1-e$), the report, θ, provides no information about the firm's true type; however, auditors are unaware that an audit failure has occurred and report θ under a mistaken belief that it is in fact the firm's true type.[16] Combining these two possibilities, the expected report for a type-ω firm going to auditor J who selects an input level e is

$$e\omega + (1 - e)(1/2).$$

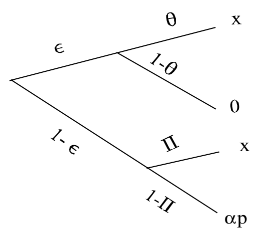
The next step involves the determination of a report-contingent market price. Because the actual value of e is unobservable, investor expectations regarding future cash flows are based on their *beliefs* ε_J regarding the level of audit accuracy. So, on observing a report θ from auditor J, investors' expectation is that it identifies the true type of the firm with probability ε_J and is "pure noise" with probability $(1-\varepsilon_J)$. When the report arises from a flawed audit, it conveys no information and the expected type *beyond the fact that it was issued by auditor J*. Therefore, the expected value contingent on audit error and report θ is $E(T|\Sigma_J) = \Pi_J$. Consequently, the expected terminal value for the firm is:[17]

$$E[\omega|\theta, \varepsilon_J, \Pi_J]\, x \;=\; \varepsilon_J \theta x + (1 - \varepsilon_J)\Pi_J x.$$

Note that the expected cash flow is a convex combination of the value, Π_J, which denotes investor's *beliefs* regarding the average probability of success across all types of firms that choose auditor J and the actual report, θ, where the weights arise from investor perceptions regarding audit accuracy.

The price paid for a firm that reports θ must also take potential recoveries through litigation into account. To analyze the role of recoveries, consider a firm that trades at a price "p" based on a report θ where p lies between the two possible realizations, 0 and x (this is always true in equilibrium as we shall demonstrate later). Then investors do not suffer any damages if x is realized but suffer damages p if 0 is realized. We simplify the computation by assuming that if a market loss of p is suffered by investors, there is an expected recovery $\alpha_J p$ from auditor J *provided that there is an error* in the audit report;[18] consequently, the expected liability for auditor J choosing input level e conditional on the firm's failure is: $\mathscr{L}_J(e, p) = \alpha_J(1 - e)p$. Notice that this formulation satisfies the fundamental requirements that the expected loss is decreasing in the level of inputs, e, while allowing the "deep pockets" to be captured through an assumption that $\alpha_B > \alpha_S$. The combined expected cash flows (provided $p \leq x$) are summarized in the next diagram (where we drop the subscript J for notational convenience).

The diagram may be understood as follows. When the auditor issues the report θ and is believed to have chosen accuracy level ε_J, the true type of the firm, ω, is expected to be the same as θ with probability ε_J— the uppermost branch of the diagram represents this event. Further, if the audit report correctly identifies the firm's type, no recoveries are possible even if the firm goes bankrupt leading to the outcome where cash flows are zero. If, on the other hand, the audit report was



wrong (lower branch), then the firm is a random member of auditor J's clientele and the expected probability of success under these circumstances is Π_J. Recoveries are impossible if the firm is successful but when the audit report is in error and the firm is unsuccessful, an event with probability $(1-\varepsilon_J)(1-\Pi_J)$, investors will recover $\alpha_J p$. Based on these potential recoveries, we now calculate the market equilibrium price.

2.1. Equilibrium Market Price

Denote the market price on observing a report θ from auditor J who is believed to have set an input level ε_J and attracted clients of average failure risk Π_J, by $p_J(\theta|\Pi_J, \varepsilon_J)$. Given the cash flow diagram presented above, the equilibrium market price, $p_J(\theta|\Pi_J, \varepsilon_J)$ solves:

$$p_J(\theta|\Pi_J, \varepsilon_J) = \varepsilon_J \theta x + (1 - \varepsilon_J)[\Pi_J x + (1 - \Pi_J)\alpha_J p_J(\theta|\Pi_J, \varepsilon_J)]. \tag{1}$$

The cash flows on the right hand side of equation (1) are the sum of expected cash flows from the firm, $\varepsilon_J \theta x + (1 - \varepsilon_J)\Pi_J x$, and expected recoveries, $\alpha_J(1 - \varepsilon_J)(1 - \Pi_J)p_J(\theta|\Pi_J, \varepsilon_J)$. Solving equation (1) for p_J,

$$p_J(\theta) = p_J(\theta|\Pi_J, \varepsilon_J) = \frac{[\varepsilon_J\theta + (1-\varepsilon_J)\Pi_J]x}{[1-(1-\varepsilon_J)(1-\Pi_J)\alpha_J]}. \qquad (2)$$

Notice that the numerator of equation (2) is at most $[\varepsilon_J + (1-\varepsilon_J)\Pi_J]x$ whereas the denominator is at least $1-(1-\varepsilon_J)(1-\Pi_J) = [\varepsilon_J + (1-\varepsilon_J)\Pi_J]$ (because $\alpha_J < 1$). Hence, the equilibrium market price, $p_J(\theta|\Pi_J, \varepsilon_J)$, is always less than the realized cash flow x (whenever the firm is successful).

2.2. *Optimal Auditor Choice*

Suppose now that each client-firm has an unbiased estimate ϕ regarding its true type, ω, i.e., $E[\omega|\phi] = \phi$, where ϕ is also distributed uniformly over (0,1). In addition, we assume that $f(\omega|\phi) > 0$ for every ω, ϕ in (0,1). Notice that the expected payoff of a firm of type ϕ is $E[\omega|\phi]x = \phi x$ and hence, a firm of higher type ϕ has greater expected payoff.

Under beliefs that the audit accuracy will be ε_J, a type ϕ-firm (whose true probability of success has density $f(\omega|\phi)$) has to consider the distribution of reports (and associated prices) it will obtain from auditor J. To arrive at this distribution, let $f(\omega|\Sigma_J)$ denote the distribution of types with auditor J. On selecting auditor J, the expected market value of a type ϕ-firm depends on two possible relationships between reported and actual types:

(1) A correct report $\theta = \omega$ and associated price $p(\theta|\Pi_J, \varepsilon_J)$ with probability $\varepsilon_J f(\theta|\phi)$,[19]

(2) A random report θ uniformly distributed over [0,1] and associated price $p(\theta|\Pi_J, \varepsilon_J)$ with probability $(1-\varepsilon_J)$.

In the first contingency where the audit process is accurate (1 above), the fact that ϕ is an unbiased estimate yields $E[f(\theta|\phi)] = E[f(\omega|\phi)] = \phi$. Combining with the linear structure in equation (2) results in the following price when the audit contains no errors:

$$\int p(\theta|\Pi_J, \varepsilon_J) f(\theta|\phi) d\theta = \frac{[\varepsilon_J\phi + (1-\varepsilon_J)\Pi_J]x}{[1-(1-\varepsilon_J)(1-\Pi_J)\alpha_J]}. \qquad (3)$$

Consider now the second possibility (2 above) where there is an error in the audit process. The audit report then randomly classifies firms. Consequently, the average report value is ½ x and linearity ensures that the average trading price resulting from an audit error is $p(\frac{1}{2}|\Pi_J, \varepsilon_J)$ as derived below:

$$\int p(\theta|\Pi_J, \varepsilon_J) f(\theta|J) d\theta = \frac{\left[\varepsilon_J \frac{1}{2} + (1-\varepsilon_J)\Pi_J\right]x}{[1-(1-\varepsilon_J)(1-\Pi_J)\alpha_J]} = p(1/2|\Pi_J, \varepsilon_J). \qquad (4)$$

Combining equations (3) and (4), under investor beliefs $\{\varepsilon_J, \Pi_J\}$, a firm of type ϕ expects the following payoff from choosing auditor J:

$$\begin{aligned} E[p|\phi \text{ chooses } J] &= \varepsilon_J \frac{[\varepsilon_J \phi + (1-\varepsilon_J)\Pi_J]x}{d_J} + (1-\varepsilon_J)\frac{\varepsilon_J 1/2x + (1-\varepsilon_J)\Pi_J x}{d_J} \\ &= \frac{\left[\varepsilon_J^2 \phi + (\varepsilon_J - \varepsilon_J^2)1/2 + (1-\varepsilon_J)\Pi_J\right]x}{d_J}, \end{aligned} \tag{5}$$

where $d_J = 1 - \alpha_J(1-\varepsilon_J)(1-\Pi_J) = (1-\Pi_J)\alpha_J\varepsilon_J + (1-(1-\Pi_J)\alpha_J)$.

In our risk-neutral, zero-discount-rate world, a firm of type ϕ would have a first-best benchmark price of ϕx under symmetric information. However, this benchmark price has to be adjusted in two directions (i) for audit errors and consequently (ii) for litigation recoveries. Audit errors affect the firm's price both directly when erroneous reports issued and indirectly through the adjustments investors make for the likelihood of error (even if an accurate reports has, in fact, been issued). The interaction of these two factors leads to the square term, ${\varepsilon_J}^2$, of the perceived audit report accuracy (numerator of equation 5).

In contrast, the denominator, d_J, represents a recovery-premium arising from the litigation process on each dollar of investment—for every dollar of expected return from the project, investors are prepared to pay $\$1/d_J$. To see this, note that the expected recovery (in case the project fails) on an investment of $\$1/d_J$ is $\alpha_J(1-\varepsilon_J)(1-\Pi_J) \times \$1/d_J$. Based on the identity $1 + [\alpha_J(1-\varepsilon_J)(1-\Pi_J) \times 1/d_J] = 1/d_J$, risk-neutral investors holding beliefs $\{\varepsilon_J, \Pi_J\}$, after taking litigation recoveries into account, will pay $\$1/d_J$ for every \$1 of expected value from the project.

Given equation (5), the firm of type ϕ will choose auditor B in preference to auditor S (respectively, auditor S in preference to auditor B) if their payoffs less audit fees are greater (respectively, smaller) with auditor B. In other words, using the expression derived in equation (5), the firm of type ϕ prefers to hire auditor B if and only if:

$$\begin{aligned} &\left[\tfrac{\varepsilon_B^2}{d_B} - \tfrac{\varepsilon_S^2}{d_S}\right]\phi x + \left[\left(\tfrac{\varepsilon_B - \varepsilon_B^2}{d_B}\right)p(1/2|\Pi_B, \varepsilon_B) - \left(\tfrac{\varepsilon_S - \varepsilon_S^2}{d_S}\right)p(1/2|\Pi_B, \varepsilon_S)\right]x \\ &\qquad + \left[\tfrac{(1-\varepsilon_B)\Pi_B}{d_B} - \tfrac{(1-\varepsilon_S)\Pi_S}{d_S}\right]x - [F_B F_S] \geq 0. \end{aligned}$$

Rearranging terms, a firm of success probability ϕ prefers auditor B if and only if:

$$\left(\frac{\varepsilon_B^2\phi + (\varepsilon_B - \varepsilon_B^2)1/2 + (1\varepsilon_B)\Pi_B}{d_B} - \frac{\varepsilon_S^2\phi + (\varepsilon_S - \varepsilon_S^2)1/2 + (1\varepsilon_S)\Pi_S}{d_S}\right)x - [F_B - F_S] \geq 0. \tag{6}$$

The terms ε_J, d_J, F_J and Π_J in equation (6) are all functions of beliefs and independent of the specific firm type ϕ. Therefore, the left- hand side of Equation (6) is linear in ϕ and is either increasing or decreasing depending on whether ${\varepsilon_B}^2/d_B$ is greater than ${\varepsilon_S}^2/d_S$ and there is a unique maximal value $\phi = \pi^*$ (where π^* depends on auditor's fee choice F_J and investor beliefs ε_J, Π_J) which satisfies the inequality in (6). The value π^* is the point at which the market segments with all types higher than π^* going to one auditor whereas types π^* or less go to the other. Before formalizing this result, one further detail, regarding the auditor's expected liability is required.

By specification, the expected liability of auditor J who sets quality e and takes on a firm of type ϕ is given by $\mathscr{L}_J(e, F|\phi) = \alpha_J(1-e)(1-\phi)[E(p|\text{audit error}, \phi)]$. By assumption, when the auditor errs (with probability $1-e$), the bankruptcy probability of the firm that is incorrectly assessed depends on the *posterior distribution* of types with auditor J after fees have been observed. That is, the expected liability to the auditor on the strategy choice $\{e, F\}$ and attracting an average client-type Π_J^* is given by: $\mathscr{L}_J(e, F) = \alpha_J(1-e)(1-\Pi_J^*)p(1/2|\Pi_J, \varepsilon_J)$. To emphasize this point, the market price, $p(1/2|\Pi_J, \varepsilon_J)$, depends on the prior belief Π_J held by investors (and is not affected by the choice of F_J) but the potential liability calculated by the auditor does depend on the fact that the average client type may change if audit fees are changed.

We next establish two benchmark results that form the starting point of our analysis.

Proposition 1

1. *Suppose that the recovery proportion from the auditors are identical, that is $\alpha_B = \alpha_S = \alpha$. Then there is an equilibrium in which all firms are indifferent across auditors and both auditors make zero profits in equilibrium.*

2. *Assume that some firm strictly prefers one of the auditors. Then there is a probability level, π^*, $\pi^* \leq 1$ such that all firms with success probabilities higher than π^* choose one auditor whereas those with lower success probabilities choose the other auditor. In particular, if $\alpha_B > \alpha_S$, there is strict preference for one of the auditors and π^* is defined.*

Proof: See Appendix B.2

The point of Proposition 1 (2) is that when $\alpha_B > \alpha_S$, the market divides into two segments (0, π^*) and [π^*, 1) with each segment choosing a different auditor (rather than in some arbitrary manner).[20] The basic intuition underlying this result may be understood by considering the auditor who supplies lower quality. This auditor must also offer lower fees in order to attract clients. If the fee discount is sufficient to attract a particular client-firm type π^*, it also suffices to attract all client-firms of type lower than π^*.

We emphasize that there is no reason, at least at this stage, why "good" (that is, low-failure rate) firms should be predisposed towards choosing the wealthy auditor. In fact, the greater insurance value provided by "wealthy" auditors is more valuable for high failure-risk firms and we shall eventually show that several additional (conflicting) factors, both endogenous and exogenous, have to be taken into account in deciding which auditor gets chosen by the low risk firms.

Proposition 1(2) implies that a meaningful value for π^* is defined when $\alpha_B > \alpha_S$ and we shall, unless otherwise stated, assume this inequality (one auditor has strictly bigger pockets than the other). In our subsequent analysis, it is convenient to use H and L to denote the auditor who gets the firms that are more likely to succeed (those with success probabilities $[\pi^*, 1)$) and the one who is selected by the lower tail firms (those with success probabilities $(0, \pi^*)$), respectively. Consequently, H (and thus, L) may be either B or S, and e_H, F_H and α_H may refer to either auditor. We note that as a consequence of the proof of Proposition 1 (2), we can identify in equilibrium, whether $H = B$ or S by comparing the magnitudes of ε_B^2/d_B and ε_S^2/d_S, quantities determined by investor beliefs. Therefore, the main focus of our analysis is trying to determine what types of beliefs can be sustained in meaningful equilibria.

Auditor H captures the market share $[\pi^*,1)$ whereas auditor L captures share $(0,\pi^*)$; because auditor choice is made *after audit fees are observed by client-firms,* both auditors and client-firms assess the average client-firm types to be $\Pi_H^* = (1+\pi^*)/2; \Pi_L^* = \pi^*/2$. In contrast, investors who do not observe audit fees continue to maintain their prior beliefs that the average types are Π_H and Π_L which may differ from Π_H^* and Π_L^*. In other words, when auditors change fees, they know their client-mix will change but are also aware that market beliefs regarding client-mix will not be revised. However, *in equilibrium,* rational beliefs require that:

$$\Pi_H = \Pi_H^* = (1+\pi^*)/2; \Pi_L = \Pi_L^* = \Pi^*/2; (\text{in addition to } \varepsilon_H = e_H \text{ and } \varepsilon_L = e_L) \tag{7}$$

These features allow us to state and establish a formal procedure for determining the equilibrium.

2.3. *Optimal Audit Strategies*

The analysis so far has determined the optimal strategy choices in the second-stage Stackelberg sub-game taking the auditors= strategy choices as given. Now, we determine the optimal strategies chosen by the auditors in the first-stage Nash game. We solve for these optimal strategies by setting up the first-order conditions for profit maximization by auditors taking into consideration the market shares that obtain in the second-stage for given beliefs $\{\varepsilon_L, \Pi_L, \varepsilon_H \Pi_H\}$.

The variable π^* provides the link between the second-stage game and the first-stage game. The prespecified beliefs together with the strategic audit choices $\{e_L, F_L, e_H, F_H\}$ determine the distribution of market prices in the second-stage (Stackelberg) game and, thereby, the client-firm type π^* that is indifferent across auditors.[21] The first-order conditions (see equations 8) demonstrate that the effects of the second-stage game on the auditor's optimal choices in the first-stage are completely captured by the value of π^* (and the prespecified beliefs).

Investor's beliefs regarding audit quality and average client types, $\{\varepsilon_L^i, \varepsilon_H^i, \Pi_H^i, \Pi_L^i\}$, are priors and independent of the choices e_L, F_L, e_H, F_H (because investors never observe these choices prior to trade). Prices are determined based on prior beliefs and the observed report, θ. Therefore, in calculating the first-order conditions in the Nash game between auditors, we can treat investor's beliefs as constants (with regard to changes in choices of e and F) and capture client-firm reactions through changes in π^*, that is, through changes in client-firm beliefs about market shares.

The expected auditor payoffs, denoted by $\mathscr{P}_L(e_L, F_L)$ and $\mathscr{P}_H(e_H, F_H)$ respectively, are given by

$$\begin{aligned}\mathscr{P}_H(e_H, F_H) &= (1-\pi^*)[F_H - C(e_H)] - \alpha_H\, p_H(1/2)x\left\{\int_{\pi^*}^{1}(1-e_H)(1-\phi)d\phi\right\}\\ &= (1-\pi^*)\left[F_H - C(e_H) - \alpha_H p_H(1/2)\left\{\left(\frac{(1-e_H)(1-\pi^*)}{2}\right)\right\}\right]\\ &= (1-\pi^*)[F_H - C(e_H) - \mathscr{L}_H(e_H, F_H)],\end{aligned}$$

$$\begin{aligned}\mathscr{P}_L(e_L, F_L) &= \pi^*[F_L - C(e_L)] - \alpha_L p_L(1/2)\left\{\int_0^{\pi^*}(1-e_L)(1-\phi)d\phi\right\}\\ &= \pi^*\left[F_L - C(e_L) - \alpha_L p_L(1/2)\left\{(1-e_L)\left(1-\frac{\pi^*}{2}\right)\right\}\right]\\ &= \pi^*[F_L - C(e_L) - \mathscr{L}_L(e_L, F_L)].\end{aligned}$$

The associated first-order conditions are:

$$\begin{aligned}\frac{\partial\mathscr{P}_H}{\partial F_H} &= [F_H - C(e_H) - \mathscr{L}_H(e_H, F_H)]\left(-\frac{\partial\pi^*}{\partial F_H}\right)\\ &\quad + (1-\pi^*)\left(\frac{\partial\mathscr{L}_H}{\partial\pi^*}\right)\left(-\frac{\partial\pi^*}{\partial F_H}\right) + (1-\pi^*) = 0,\\ \frac{\partial\mathscr{L}_L}{\partial F_L} &= [F_L - C(e_L) - \mathscr{L}_L(e_L, F_L)]\left(\frac{\partial\pi^*}{\partial F_L}\right) + \pi^*\left(\frac{\partial\mathscr{L}_L}{\partial\pi^*}\right)\left(\frac{\partial\pi^*}{\partial F_L}\right) + \pi^* = 0, \qquad (8)\\ \frac{\partial\mathscr{P}_H}{\partial e_H} &= (1-\pi^*)\left[-\frac{\partial C(e_H)}{\partial e_H} - \frac{\partial\mathscr{L}_H(e_H, F_H)}{\partial e_H}\right] = 0,\\ \frac{\partial\mathscr{P}_L}{\partial e_L} &= \pi^*\left[-\frac{\partial C(e_L)}{\partial e_L} - \frac{\partial\mathscr{L}_L(e_L, F_L)}{\partial e_L}\right] = 0.\end{aligned}$$

In understanding the first-order conditions, recall that $\mathscr{L}_J(e_J, F_J) = \alpha_J(1-e)(1-\Pi_J^*)p(1/2|\Pi_J, \varepsilon_J)$ is a function of e_J and indirectly (through Π_J^*) a function of both π_J^* and F_J. We have to solve for values of e_L, F_L, e_H, F_H that satisfy these equations for the given values of $\{\pi^*, \Pi_H^i, \Pi_L^i, \varepsilon_H^i, \varepsilon_L^i\}$ and verify the following two conditions:

(1) The optimal responses $\{e_L, e_H\}$ are the same as the prior beliefs $\{\varepsilon_L^i, \varepsilon_H^i\}$,[22] and

(2) The optimal responses $\{F_L, F_H\}$ lead to indifference across auditors for firm-type π^* that makes the quantities $\{\pi^*/2, (1+\pi^*)/2\}$ agree with the prior beliefs, $\Pi_L^i \Pi_H^i$.

A key feature of the first-order conditions is that it conforms with a structure where client-firms observe fee choices but not audit quality choices, whereas investors observe neither audit fees nor audit quality at the time of trade. Therefore, neither investors nor client-firms update beliefs regarding audit quality whereas client-firms revise their beliefs about the auditor's clientele based on the fees observed by them.[23] The solution of the first-order conditions (8) formally determines a sequential equilibrium under a specification that client-firms revise their beliefs regarding auditor's clientele but not about auditor quality when they observe out-of-equilibrium audit fees.

From equation (6) we obtain (irrespective of which auditor is H) that:

$$\left[\frac{\varepsilon_H^2}{d_H} - \frac{\varepsilon_L^2}{d_L}\right]\pi^* x + \left[\left(\frac{\varepsilon_H - \varepsilon_H^2}{d_H}\right)\frac{1}{2} - \left(\frac{\varepsilon_L - \varepsilon_L^2}{d_L}\right)\frac{1}{2}\right]x + \left[\frac{(1-\varepsilon_H)\Pi_H}{d_H} - \frac{(1-\varepsilon_L)\Pi_L}{d_L}\right]x = [F_H - F_L]$$

This equation may be simplified as:

$$\left[\frac{\varepsilon_H^2}{d_H} - \frac{\varepsilon_L^2}{d_L}\right]\left[\pi^* - \frac{1}{2}\right]x + \left[\frac{\varepsilon_H}{d_H} - \frac{\varepsilon_L}{d_L}\right]\frac{x}{2} + \left[\frac{(1-\varepsilon_H)\Pi_H}{d_H} - \frac{(1-\varepsilon_L)\Pi_L}{d_L}\right]x = [F_H - F_L] \tag{9}$$

In equation (9), π^*, F_H and F_L are endogenous choices whereas all other terms are functions of the prior beliefs. Because $\varepsilon_H^2/d_H > \varepsilon_L^2/d_L$, we find that π^* increases as the spread of audit fees, $F_H - F_L$, increases. In other words, the market share of the low auditor, L, represented by $(0,\pi^*)$, increases when the fee difference between the two auditors widens. Differentiating the expression in equation (9) (with respect to F_H and F_L , respectively) yields:

$$\left[\left(\frac{\varepsilon_H^2}{d_H}\right) - \left(\frac{{\varepsilon_L}^2}{d_L}\right)\right]\frac{\partial \pi^*}{\partial F_H} = \frac{1}{x}; \qquad \left[\left(\frac{\varepsilon_H^2}{d_H}\right) - \left(\frac{{\varepsilon_L}^2}{d_L}\right)\right]\frac{\partial \pi^*}{\partial F_L} = -\left[\frac{1}{x}\right]. \tag{10}$$

Equation (10) shows that the elasticity of market share in price is identical across the auditors. Substituting the expressions for the derivatives of π^* derived in equation (10), the system of equations (8) and (9) consists of five equations in the five unknowns F_H, e_H, F_L, e_L and π^*. To solve this system, we treat π^* as a free variable in the system of first-order conditions (8) and derive a solution $\{e_S(\pi^*), F_S(\pi^*), e_B(\pi^*), F_B(\pi^*)\}$ for the other variables in terms of π^*. Substituting these expressions back, the left-hand side of equation (9) (with $\varepsilon_J = e_J(\pi^*)$) may be solved as a function of π^* and root yields a potential equilibrium (see Appendix C for an implementation of this procedure). While this technique will identify an equilibrium if it exists, there are some technical difficulties regarding the existence of an equilibrium in our context (see Appendix A). However, when one auditor is "much wealthier'' than the other, existence may be easily demonstrated as described in Proposition 2.

Proposition 2

Suppose that only the big auditor can pay damages, that is, $\alpha_B > 0$ and $\alpha_S = 0$. *Then: (i) an equilibrium always exists and (ii) in any equilibrium, the wealthy auditor is always chosen by the high value firms, i.e,* $B = H$.

Proof:

If $\alpha_S = 0$, then it follows immediately from equation (8) that $C'(e_S) = 0$. However, if $C(e)$ is increasing, this means that the corner solution $e_S = 0$ is the only possibility. (When $\pi^* = 0$, every choice of e_S including $e_S = 0$ is optimal). Further, because $\alpha_B > 0, C(e_B) > 0$ and hence, $e_B > 0$, that is, the best responses for the auditors is to set $e_B > e_S = 0$. In this circumstance, the wealthy auditor always captures the high end of the market, that is, $\Pi_B > \Pi_S$ and this fact ensures the existence of an equilibrium (see Appendix A and also the example in Appendix C).

Proposition 2 sets up a benchmark where one auditor is ''very'' wealthy relative to the other. This is in polar contrast to Proposition 1 (1) where both auditors are equally wealthy. The second part of the proposition that high-risk clients choose the auditor with shallow pockets requires some elaboration. Indeed, client-firm's of high failure risk benefit more from the protection arising from ''big pockets'' and may be expected to select the wealthy auditor. The economic factor underlying auditor selection in Proposition 2 is that small auditors, because they pay low damages, can compete more aggressively on fees. Wealthy auditors, facing stiff damages, either have to provide more accurate or charge much higher fees. In either case, they lose the high-risk client-firms who dislike both accurate audits and higher fees.

The next proposition provides some interesting information about the characteristics of equilibria within our setting. Specifically, we can now demonstrate that high audit quality is needed to attract high value firms.

Proposition 3

In equilibrium, the auditor who attracts high-type firms supplies better quality and thus reports more accurately. If the auditor with shallow pockets is to attract the high-type firms, he has to provide strictly greater quality than the wealthy auditor.

Proof:

Setting $\Pi_H = (1+\pi^*)/2$ *and* $\Pi_L = \pi^*/2$ *in the first-order conditions with regard to e* (last two Equations of (8)), we get:

$$\begin{array}{c} -[C'(e_H) + \mathscr{L}_H(e_H, F_H)] = 0 \rightarrow C'(e_H) = \alpha_H(1-\Pi_H)\ p_H(1/2) \\ -[C'(e_L) + \mathscr{L}_H(e_H, F_H)] = 0 \rightarrow C'(e_L) = \alpha_L(1-\Pi_L)\ p_L(1/2) \end{array}. \tag{11}$$

Auditor H gets the market share $[\pi^*,1)$ so $e_H{}^2/d_H \geq e_L{}^2/d_L$ (Proposition 1). Suppose that, in addition, we also have $e_H < e_L$ (we shall shortly prove that this situation results in a contradiction). Then: (i) $\alpha_H(1-\Pi_H) \geq \alpha_L(1-\Pi_L)$ and (ii) $d_H < d_L$

(Lemma (part 3)). From these relations and the fact that $\Pi_H > \Pi_L$ Equations (11) imply that $C'(e_H) > C'(e_L)$. But this relation and the fact that $C'(e)$ is increasing in e force $e_H > e_L$(recall that $C(e)$ is convex) , a contradiction of the hypothesis that $e_H < e_L$. Therefore, we have shown that $e_H \geq e_L$, that is, the auditor who captures the high end of the pool also supplies more accurate reports.

If $e_B = e_S$, and S captures the high-end pool, then $\Pi_B > \Pi_S$ but the fact that $\alpha_B > \alpha_S$ implies that $d_B < d_S$ and hence that $e_B^2/d_B \geq e_S^2/d_s$ implying that the big auditor captures the high end pool, a contradiction. Thus, if $e_B = e_S$, B must capture the high-type firms.

It seems "obvious" that Proposition 3 should hold, that is, as a consequence of adverse selection, better firms should prefer greater accuracy in reports. The reason that a proof is needed (and its relative complexity) stems from the rational expectations framework. Formally, it is possible that prior beliefs that good firms select the auditor who supplies low quality and sets low fees will result in prices that sustain those beliefs. In essence, Proposition 3 shows that prior beliefs that high quality client-firms select low-quality low-fee auditors cannot be sustained in equilibrium.

The previous result shows that, in equilibrium, the auditor who captures the high end of the market in equilibrium will also supply more accurate reports. The combination of clients of lower failure risk together with the greater accuracy of the report necessarily leads to a lower litigation rate. Nevertheless, if wealthy auditors capture the high end of the market, in some parameter ranges (such as the one identified below in Proposition 4), they could end up paying larger amounts to settle their cases. The interesting inference from this result is that the level of damages may not provide a suitable empirical proxy for audit quality, that is, it may be incorrect to conclude that auditors who are penalized more heavily for audit failures are providing lower quality audits.

Proposition 4

If the cost function for accuracy is such that $(1-e)C'(e)$ is increasing in e, then auditor H pays a greater (expected) amount in equilibrium (per client) to settle claims arising from audit failures than auditor L.

Proof:

Multiplying the first-order conditions for optimal quality, in equation (11), by $(1-e_J)$ we obtain:

$$\begin{aligned} (1-e_H)C'(e_H) &= \left(\tfrac{\alpha_H(1-e_H)(1-\Pi_H)}{dH}\right)p_H(1/2) = \mathscr{L}_H(e_H, F_H) \\ (1-e_L)C'(e_L) &= \Pi_L\left(\tfrac{\alpha_L(1-e_L)(1-\Pi_L)}{d_L}\right)p_L(1/2) = \mathscr{L}_L(e_L, F_L) \end{aligned} . \tag{12}$$

Because the auditor H attracts better firms, $e_H > e_L$ (Proposition 3). The assumption of this proposition is that $(1-e)C'(e)$ is increasing in e and it follows that $(1-e_H)C'(e_H) > (1-e_L)C'(e_L)$ and by equation (12), that $\mathscr{L}_H(e_H, F_H) > \mathscr{L}_L(e_L, F_L)$;

that is, the average loss per client for auditor H, which is, by definition, $\mathscr{L}_H(e_H, F_H)$, exceeds that of auditor L, $\mathscr{L}_L(e_L, F_L)$.

We stress that the result of Proposition 4 is driven by the fact that lawsuits related to high-type firms involve greater losses for the auditors than those related to low-type firms. This factor may outweigh the lower litigation risk under some parametric conditions.

The next proposition, our main result, shows why ''boutique auditors'' may not be observed in equilibrium. By this, we mean an equilibrium where one or the other auditor has market share consisting of a small group of high type clients. We show that as long as the cost of reducing errors in the audit report exceeds a certain threshold, the auditor who captures the high-type client also is dominant in terms of market share. In addition, this result also indicates why wealthy auditor B typically captures a large market share. The higher litigation burden induces auditor B to supply more accurate reports than auditor S. Therefore, by Proposition 3, they attract the high-type (low-risk) firms and as a consequence of Proposition 5, capture a large share of the market.

The economic motivation underlying Proposition 5 is that as the average client-type with auditor H increases (i.e., π^* increases), the marginal litigation cost drops as a function of audit accuracy, e. However, as observed in the arguments leading to Proposition 4, in equilibrium, the marginal cost of audit accuracy, $C'(e)$, is the same as the marginal cost of litigation. Therefore, at high values of π^*, when the marginal cost of litigation for auditor H is low, that auditor has little or no incentives to supply accurate audits; however, as audit quality, e, drops, the high-type clients will switch to the other auditor. This argument demonstrates that equilibrium with high levels of both π^* and e are untenable.

Notice that the assumption of audit quality being supplied under moral hazard also plays an important role in this argument. In our framework of moral hazard, auditors cannot precommit to (high) levels of audit quality that turn out to be *ex post* uneconomical. So audit firms that attract only high-type clients will not perform expensive audits.

Some restrictions have to be placed on audit costs in order to make this intuition rigorous. If the cost of litigation is very high relative to the cost of delivering accurate audits, auditors may prefer to supply very high quality even though the litigation probability has dropped close to zero. In such extreme circumstances, boutique auditors may be present in equilibrium. But as long as the marginal cost of increasing audit accuracy is not too low, high values of π^* imply low values of e.[24] As a consequence, the optimal strategy is to go for greater market share even if this increases the expected liability payments. The larger market share and attendant increase in litigation risk helps to convince investors and client-firms that the auditor will maintain a greater level of accuracy. In turn, the perception of higher quality allows the auditor to attract low-risk clients and increase profits.

Proposition 5

The auditor who attracts the high-type clients also captures more than half the market and obtains greater total profits provided that the marginal cost of reducing errors,

$C'(e)$, is at least as large as $(\alpha_B/2)e\,x$. In addition, if $\alpha_S = 0$ (as in Proposition 2), the wealthy auditor captures larger market share, that is, B is large both in terms of wealth and in terms of market share.

Proof: See Appendix B.3

Proposition 5 establishes two points: (i) that ''boutiques'' are not tenable in equilibrium and (ii) that the auditor who has big pockets typically captures a large market share. The first point follows directly from the result that the auditor who attracts high-type firms also captures more than half the market share precluding the possibility of a high quality auditor with small market share. The second point is not as straightforward. When $\alpha_S = 0$, the wealthy auditor, *B*, always implements a higher level of *e*, gets the high-end of the market (Propositions 2), and, hence, by Proposition 5, also gets the major share of the market. That is, auditor *B* is ''Big'' both in terms of pockets and in terms of market share. The same argument holds whenever differences in the wealth levels are sufficiently large. With sufficient wealth difference, the belief that the ''wealthy auditor will capture a large market-share'' is a self-fulfilling phenomenon that can be sustained in at least one equilibrium (using the technique outlined in Appendix A) whereas beliefs that the auditor with the big pockets attracts low-type clients and small market share are typically not sustainable (and definitely unsustainable when $\alpha_S = 0$).

Another interesting and significant feature of the equilibrium is that both auditors will make positive profits but the auditor with the larger market share will make greater profits. The result suggests one reason why the ''Big Four'' may be more profitable than smaller auditors. In addition, this result shows that the situation where the wealthy auditor captures large market share is ''stable'' in the sense that they will maintain or increase the differences in wealth levels and continue to hold a competitive edge over the less wealthy auditor. The example given in Appendix C demonstrates how equilibrium can be derived in this sequential setting and illustrates all the issues raised in Proposition 5—in particular, auditor *B* is ''big'' both in terms of pockets and in terms of market share and attains greater profitability.

We now consider a situation where both auditors are constrained, by regulation, to set very high minimum levels of audit quality. In particular, if any level higher than this statutory minimum is uneconomical, we show that the competitive power of the wealthy auditor is increased and that he captures at least two-thirds of the market in equilibrium. This result is established without any reference to the cost function (as in Proposition 5) because the level of audit inputs is driven by statute rather than endogenously through the characteristics of client-firms.

2.4. *Standard Setting and Audit Industry Structure*

In this section, we examine a situation where the regulatory structure forces both auditors to choose the same investment in audit quality. While this is undoubtedly a special case, it arises fairly plausibly under negligence rules. Suppose that the auditor faces a punitive fine P whenever quality drops below a certain minimum level $e^* < 1$, that is, the liability rule, β, is given by:[25]

$$\beta(e) = \begin{cases} -P & \text{if } e < e^* \\ 0 & \text{if } e \geq e^*. \end{cases}$$

If P is sufficiently large so as to deter the choice of quality $e < e^*$ and that $C'(e^*)$ is so high that quality levels greater than e^* are suboptimal for either auditor. Then both auditors will set exactly the same level of quality $e = e^*$.[26] We show that such a penalty structure that forces both auditors to set the same prescribed quality level e^* results in greater market share for the auditor with bigger pockets.

Proposition 6

Suppose that the minimum standard $e^* > 0$ is so high that neither auditor will choose input levels higher than e^*. Then the auditor with deeper pockets captures at least two-thirds of the market.

Proof: See Appendix B.4.

When both auditors choose the same levels of "e" (because of statutory considerations), the highest-type clients choose the auditor with deep pockets and the optimal profit strategy for the wealthy auditor is to capture a large market share (as in Proposition 5). However, when both auditors are forced to report with the same level of accuracy, the small auditor can no longer lure away high failure-risk firms by offering them lower quality at a reduced fee (because a minimum accuracy levels are statutorily enforced) or the low failure-risk firm by offering greater quality (because the cost of such a high level of inputs is prohibitively expensive). Therefore, the advantage of the wealthy auditor is enhanced and they capture at least two-thirds of the market.

3. Conclusion

We analyze a situation where two auditors, B (Big Pockets) and S (Shallow Pockets) compete for a set of clients with differing risk profiles. The auditors are otherwise identical and are distinguished solely by the fact that B has bigger pockets relative to S and a greater incentive to avoid litigation. Our formal analysis demonstrates that if the marginal cost of generating greater accuracy is significant relative to the potential liability, boutique auditors (i.e., audit firms with a small base of high-quality clients) will not be viable in equilibrium. In addition, if market perceptions are that auditors with bigger pockets will supply higher quality audits, any equilibrium conforming with these beliefs will result in the wealthy auditor capturing a major share of the market and attaining a greater level of profitability. In other words, our results also demonstrate how initial wealth differences that may arise randomly lead to an auditor who dominates the market and maintains these initial random wealth differences. This equilibrium is descriptively accurate in terms of the current market structure where almost all of the Fortune 500 firms are audited by the "Big Four".

We draw several additional inferences concerning fee structures, audit quality, litigation rates and the effects of regulation. The more accurate auditor attracts better client-firms. However, if wealthier auditors are thought to be attracting higher quality firms, low quality firms might be tempted to select these auditors in order to capitalize on a "halo" effect. Both high fees and more accurate audits may function as mechanisms for deterring high-risk firms as clients. That is, observed fee differentials between "Big Four" auditors and smaller auditors may partially be driven by the strategic need to deter low-quality clients who may wish to be associated with the portfolio of good clients that typically choose to be audited by the "Big Four." As a further consequence, auditors who supply lower accuracy compete more effectively for lower-type clients. Regulations that force high audit accuracy on both types of auditors blunt this competitive edge and favor the wealthy auditor.

Our analysis assumes that all auditors have equally efficient technologies for producing quality reports. Differences in audit technology may provide an alternative explanation for some of our results regarding market shares. However, we emphasize that introducing differential cost functions will not alter the economics underlying our main results. Specifically, if the big auditor has a cost advantage (in addition to deeper pockets), they gain extra market share, but neither the nature of the segmentation of the market nor the fact that the small auditor can compete more effectively for "low" value client-firms is affected by introducing differential audit costs.

Our last implication concerns the correspondence between observed litigation data and the accuracy of information supplied in equilibrium. Given that good firms are less likely to fail and provoke litigation against auditors, systematic selection of one auditor by all good firms "biases" the observed litigation rate. Therefore, it is possible to observe differential litigation rates against auditors even when both auditors supply reports of equal accuracy. However, good firms are also likely to trade at higher prices. Therefore, in the event of their failure, the level of market losses, and hence damage claims, are substantially higher. In turn, this might lead to greater per-client losses for the auditor who attracts better quality firms through more accurate reporting policies. In particular, wealthier auditors may attract higher quality firms in equilibrium and end up appearing inefficient (when measured through the fraction of their fees used in litigation settlements) even though they supply more accurate reports.

Appendix A: The Equilibrium Process

Formally, an equilibrium (in pure strategies) for our model consists of a specification of the quantities:

$$\{\{e_B, F_B\}\{e_S, F_S\}; \quad \{\varepsilon_B^i, \Pi_B{}^i, \varepsilon_S{}^i \Pi_S^i\}; \qquad p_B(\theta|\varepsilon_B{}^i, \Pi_B{}^i); \qquad p_S(\theta|\varepsilon_S{}^i, \Pi_S{}^i)\}$$

where $\{\varepsilon_J^i, \Pi_J^I\}$ and $\{\varepsilon_J^f, \Pi_J^f\}$ denote beliefs held by investors and firms regarding audit quality and the average firm type selecting auditor J. Market price schedules p_B

and p_S satisfy equation (2) (given beliefs $\{\varepsilon_J^i, \Pi_J^I\}$); the choice of auditor by each firm maximizes expected payoffs as described in equation (6) (given beliefs $\{\varepsilon_J^f, \Pi_J^f\}$); the auditors' strategies should be optimal given the expectations and consequent prices offered by investors; last, beliefs are rational, that is, $\varepsilon_J^i = e_J = \varepsilon_J^f$, while Π_j^i, Π_J^f, and π^* satisfy equations (7) (appropriately depending on whether $H = B$ or S).

Suppose that a set of strategies $\{e_B, F_B, e_S, F_S\}$ and beliefs $\{\varepsilon_J^f = \varepsilon_J^i = e_J, \Pi_J^i = \Pi_J^f = \Pi_J\}$ ($J = B$ or S) are given with ε_J, e_J in [0,1], F_J in [0, M] (where M exceeds the cost of going private and is a fee level that is unacceptable for any firm) and Π_B, Π_S are such that $|\Pi_B - \Pi_S| = 1/2$. We may compute the best responses $\{e_B^r, F_B^r, e_S^r, F_S^r, \Pi_B^r, \Pi_S^r\}$ by:

(i) calculating d_J and determining the auditor who captures the high value firms through equation (6) and determining $\pi^{*\mathrm{r}}$;

(ii) solving the first-order conditions (8);

(iii) solving the optimal auditor choice equation (9).

where the values Π_J^r are determined by π^{*r} and the auditor who captured the high end of the market under step (i) above. Notice that $|\Pi_B^r - \Pi_S^r| = 1/2$ by definition.

In other words, we have constructed a best response correspondence that maps ▯′→′▯where▯ is given by:

$$▯=[0,1]\times[0,M]\times[0,1]\times[0,M]\times\left\{(\Pi_B-\Pi_S)|0\leq\Pi_B\leq 1, 0\leq\Pi_s\leq 1, |\Pi_B-\Pi_S|=1/2\right\}.$$

Unfortunately, the set ▯ is not convex (this is always a problem with pure strategies), and therefore, the best response correspondence may not have a fixed point. Fortunately, an intuitive economic restriction does guarantee the existence of an equilibrium. Suppose that when investors believe the big auditor works harder ($\varepsilon_B^i \geq \varepsilon_S^i$) and attracts the high end of the market ($\Pi_B = \Pi_S + 1/2$), the best response for the small auditor is indeed to concede this market share to the big auditor. Then the best response correspondence maps the set:

$$▯^c=[0,1]\times[0,M]\times[0,1]\times[0,M]\times\left\{(\Pi_B-\Pi_S)|0\leq\Pi_B\leq 1, 0\leq\Pi_S\leq 1, (\Pi_B-\Pi_S)=1/2\right\}$$

into itself. Further, ▯^{c} is convex (and compact). In consequence, the best response function does have a fixed point, the fixed point constitutes an equilibrium and in this equilibrium, $\Pi_B = 1/2 + \Pi_S$, that is, the big auditor captures the high end of the market.

One last technical point must be mentioned: if $\pi^* = 0$ or 1, that is, if one auditor captures the whole market, then the beliefs regarding the average type choosing the other auditor is not well-specified. We shall impose the restriction that in these limit cases, the equation $\Pi_B - \Pi_S = 1/2$ continues to hold.

Appendix B: Technical Proofs

Appendix B.1: Lemma

The lemma given below summarizes some relationships that will be used in subsequent proofs.

Lemma

(1) e/d_J and e^2/d_J are increasing in e (with d_J treated as a function of e);

(2) $\varepsilon_J^2/d_J < \varepsilon_J$ for $J = B,S$;

(3) Let ε_H ,ε_L, denote any two values such that $\varepsilon_L^2/d_L < \varepsilon_H^2/d_H$. If , in addition, $\varepsilon_H < \varepsilon_L$, then $\alpha_H(1-\Pi_H) > \alpha_L(1-\Pi_L)$ and $d_H < d_L$.

Proof:

(1) Noting that $d_J = 1 - \alpha_J(1 - \Pi_J)(1 - e)$ is a function of e, direct differentiation shows that:

$$\frac{\partial}{\partial e}\left(\frac{e}{d_J}\right) = \frac{(1-\alpha_J(1-\Pi_J)(1-e)) - e\alpha_J(1-\Pi_J)}{d_J^2}$$
$$= \frac{1-\alpha_J(1-\Pi_J)}{d_J^2} > 0$$

where the last inequality follows from the fact that e, α_J and Π_J are all less than 1. Hence, e/d_J is strictly increasing for positive e. Because e/d_J is increasing and positive, $e^2/d_J = e \times (e/d_J)$ is the product of two positive increasing functions and hence positive and increasing.

(2) By direct evaluation, we find that when $e = 1$, $e/d_J(e) = 1$. It follows from (1) that for $0 \le e \le 1$, $e/d_J(e) \le 1$, implying $e^2/d_J(e) \le e$. The result follows on evaluating at $e = \varepsilon_J$.

(3) Now suppose that $\varepsilon_H < \varepsilon_L$ and $\varepsilon_H{}^2/d_H\ \varepsilon_L{}^2/d_L$.

$$\frac{\varepsilon_L^2}{1-\alpha_L(1-\varepsilon_L)(1-\Pi_L)} = \frac{\varepsilon_L^2}{d_L} \le \frac{\varepsilon_H^2}{d_H} = \frac{\varepsilon_H^2}{1-\alpha_H(1-\varepsilon_H)(1-\Pi_H)}$$
$$< \frac{\varepsilon_L^2}{1-\alpha_H(1-\varepsilon_L)(1-\Pi_H)}.$$

where the strict inequality is a consequence of the assumption that $\varepsilon_H < \varepsilon_L$ and the first part of the lemma. But comparing the first and the last terms shows that $\alpha_H(1-\Pi_H) > \alpha_L(1-\Pi_L)$. Because of this inequality and the fact that $\varepsilon_H < \varepsilon_L$, it follows that: $\alpha_H(1-\Pi_H)(1-\varepsilon_H) > \alpha_L(1-\Pi_L)(1-\varepsilon_L)$ and hence, that $d_H < d_L$.

Appendix B.2: Proof of Proposition 1

All firms are indifferent across auditors only if the left-hand side of equation (6) is identically zero for every choice of ϕ. In this circumstance, we may assume that firms randomly choose auditors and that the average failure probability of firms with each auditor is $\Pi_J = ½$

(1) Let $\mathscr{L}(e)$ denote the function $\alpha(1-e)(1/2)p(1/2|\Pi_J = 1/2, e) = [\alpha(1-e)(1/2)][1/2(x/d_J)]$ and let e^* be the level of inputs minimizing the total cost $C(e) + \mathscr{L}(e)$. Let $F^* = C(e^*) + \mathscr{L}(e^*)$ the break-even fee at this level of inputs. Then : 9 8e^*,F^* ≅ 8e^*,F^* ≅; $\Pi_B = ½$; $\Pi_S = ½$; $\varepsilon_B = \varepsilon_S = e^*$ A, with associated prices constitutes an equilibrium. That is, given these strategic choices by the auditor:

(A) All firms are indifferent across auditors and therefore randomly and independently choose one or the other auditor. Thus, the average success rate of firms going to each auditor is ½.

(B) Given this policy of auditor choice by firms, both auditors set quality at e^*, and the fee charged allows each of them to break even.

(C) Any change in e increases costs to the auditor without affecting any of the other factors (because changes in e are unobservable). Any increase in (observable) fees leads to the loss of all market share whereas decreases below the break-even level F^* are not tenable. Therefore, (both) auditors are satisfied with the strategy choices $\{e^*, F^*\}$.

(D) Investors believe (accurately) that the success rate of firms going to each auditor is the same (and equal to ½) and that each auditor has chosen inputs e^*.

(2) The continuity of equation (6) in ϕ implies either (i) the inequality is satisfied for every ϕ or (ii) it fails for every ϕ or (iii) there is a value, $0 < \pi^* < 1$, solving equation (6). The *linearity* of the price in ϕ (see equation (6)), implies that depending on whether $\varepsilon_B^2/d_B > \varepsilon_S^2/d_S$ (respectively, $\varepsilon_B^2/d_B < \varepsilon_S^2/d_S$), all firms above the cutoff π^* go to auditor B (respectively, to auditor S). If $\varepsilon_B{}^2/d_B = \varepsilon_S{}^2/d_S$, then ϕ drops out of Equation (6) and the left-side of this equation has the same sign for all ϕ, that is, all firms choose the same auditor unless if the equation is identically zero.

Suppose now that $\alpha_B > \alpha_S$ and that all firms are indifferent across auditors, that is, the left-hand side of equation (6) is zero for every ϕ; we show that such a situation is untenable in equilibrium. First, under indifference, $\Pi_B = ½ = \Pi_S$ (as in Part (A) above) and we have $d_B = 1 - \alpha_B (1 - \varepsilon_B)$ ½ ; $d_S = 1 - \alpha_S (1 - \varepsilon_S)$ ½. Next the left-hand side of equation (6) is zero for every ϕ only if $\varepsilon_B{}^2/d_B = \varepsilon_S{}^2/d_S$. We show that this combination results in a contradiction. If $\varepsilon_B \geq \varepsilon_S$ then by the first part of the lemma:

$$\frac{\varepsilon_B^2}{d_B} = \frac{\varepsilon_B^2}{1 - \alpha_B(1 - \varepsilon_B)(1/2)} \geq \frac{\varepsilon_S^2}{1 - \alpha_B(1 - \varepsilon_S)(1/2)}$$
$$> \frac{\varepsilon_S^2}{1 - \alpha_S(1 - \varepsilon_S)(1/2)} = \frac{\varepsilon_S^2}{d_S}$$

but $\varepsilon_B{}^2/d_B > \varepsilon_S{}^2/d_S$ means that the left-hand side of equation (6) cannot be identically zero for every ϕ. Therefore, $\alpha_B > \alpha_S$ and indifference for all client-firms (in equilibrium) implies $\varepsilon_B < \varepsilon_S$ and consequently that $d_B < d_S$,(for otherwise, $\varepsilon_B{}^2/d_B < \varepsilon_S{}^2/d_S$). Also, under indifference, $\Pi_B = ½ = \Pi_S$ and the expected litigation cost becomes $\mathscr{L}_J(e) = [\alpha_J(1\text{-}e_J)/4d_J]x$. From $d_B < d_S$, we infer that:

$$\mathscr{L}_B(e) = (\alpha_B x)/4d_B) > (\alpha_S x)/4d_S) = -\mathscr{L}S(e).$$

In equilibrium, auditor J chooses e_J to minimize $C(e) + \mathscr{L}_J(e)$, that is, $C'(e_J) = -\mathscr{L}'_J(e_J)$. Thus $-\mathscr{L}'_B(e) > -\mathscr{L}'_S(e) \Rightarrow C'(e_B) > C'(e_S)$, and from the convexity of $C(e)$ we may conclude that $e_B > e_S$. Because $e_J = \varepsilon_J$ for $J = B$, S, in equilibrium, we have reached a contradiction; that is, we have shown that if $\alpha_B > \alpha_S$ and $\varepsilon_B{}^2/d_B = \varepsilon_S{}^2/d_S$, then $\varepsilon_B < \varepsilon_S$ and $e_B > e_S$, which is not possible in equilibrium.

Appendix B.3: Proof of Propositions 5

Using equation (10), the first two equations of (8) may be rewritten as follows:

$$\begin{aligned}&\left[F_H - C(e_H) - \mathscr{L}_H(e_H,F_H) + (1-\pi^*)\tfrac{\partial \mathscr{L}_H}{\partial \pi^*}\right] = (1-\pi^*)\left(\tfrac{\partial \pi^*}{\partial F_H}\right)^{-1}\\&\left[F_L - C(e_L) - \mathscr{L}_L(e_L,F_L) - \pi^*\tfrac{\partial \mathscr{L}_L}{\partial \pi^*}\right] = -\pi^*\left(\tfrac{\partial \pi^*}{\partial F_L}\right)^{-1} = \pi^*\left(\tfrac{\partial \pi^*}{\partial F_H}\right)^{-1}.\end{aligned} \tag{B1}$$

Substituting from equation (B1) into the expression for auditor profit (reproduced below) yields:

$$\begin{aligned}\mathscr{P}_H(e_H,F_H) = (1-\pi^*)[F_H - C(e_H) - \mathscr{L}_H(e_H,F_H)] = (1-\pi^*)^2\left[\left(\tfrac{\partial \pi^*}{\partial F_H}\right)^{-1} - \tfrac{\partial \mathscr{L}_H}{\partial \pi^*}\right]\\ \mathscr{P}_L(e_L,F_L) = \pi^*[F_L - C(e_L) - \mathscr{L}_L(e_L,F_L)] = (\pi^*)^2\left[\left(\tfrac{\partial \pi^*}{\partial F_H}\right)^{-1} + \tfrac{\partial \mathscr{L}_L}{\partial \pi^*}\right).\end{aligned} \tag{B2}$$

Writing p_J (1/2) for $p(1/2 \mid \Pi_J, \varepsilon_J)$ and using the identity $\mathscr{L}_J(e,F) = \alpha_J(1-e)(1-|\Pi_J^*)p_J(1/2)$ we get:

$$\begin{aligned}\tfrac{\partial \mathscr{L}_H}{\partial \pi^*} &= \left[\tfrac{\alpha_H(1-e_H)}{2}\right]p_H(1/2) \leq 0,\\ \tfrac{\partial \mathscr{L}_L}{\partial \pi^*} &= \left[\tfrac{\alpha_L(1-e_L)}{2}\right]p_L(1/2) \leq 0.\end{aligned} \tag{B3}$$

We note that one would normally expect the derivative of $\mathscr{L}_J$ in π^* to be negative for the following intuitive reason: as π^* increases, the average type of client-firm increases for *both* auditors H and L, and consequently, it is reasonable to expect that $\mathscr{L}_J$ declines in π^* for both auditors.

Equations (B2) and (B3) together demonstrate that a larger market share for H, i.e., $\pi^* < 2 < (1 - \pi^*)$ implies both greater total profits and greater per-client profits for H.

The key step in establishing the proposition is that if $C'(e) \geq (\alpha_B ex)/2$, an auditor who attracts the high-type client-firms $[\pi^*, 1)$ will supply effort $e_H \leq 1/2$ whenever $\pi^* \geq 1/2$. This result reflects the economic fact that attracting higher client-types results in a reduced need for audit accuracy. From the third equation in (8) and the hypothesis that $C'(e) \geq (\alpha_B ex)/2$, we infer:

$$\frac{\alpha_B e_H x}{2} \leq \alpha_H(1 - \Pi_H)\, p_H(1/2) \;=\; C(e_H) \quad \Rightarrow \quad e_H x \leq 2(1 - \Pi_H)p_H(1/2).$$

When $\pi^* > 1/2$, $\Pi_H > 3/4$; further, prices can never exceed x . Combining these two facts we obtain that $e_H \leq 2 \times 1/4 = 1/2$. We show that the combination $e_H \leq ½$ and $\pi^* > ½$ leads to a contradiction; it then follows that $\pi^* \leq 1/2$ in equilibrium.

Subtracting the second equation in (B1) from the first yields:

$$\begin{aligned}&[F_H - F_L] - \{[C(e_H) - C(e_L)] - [\mathscr{L}_H(e_H, F_H)\mathscr{L}_L(e_L, F_L)]\\ &\qquad - \left[(1 - \pi^*)\frac{\partial \mathscr{L}_H}{\partial \pi^*} + \pi^* \frac{\partial \mathscr{L}_L}{\partial \pi^*}\right]\Bigg\} = (1 - 2\pi^*)\left[\frac{\partial \pi^*}{\partial F_H}\right]^{-1}\end{aligned} \tag{B4}$$

Next, $C(e) + \mathscr{L}_H(e, F_H)$ is minimized at $e = e_H$; so $C(e_H) + \mathscr{L}_H(e_H, F_H) \geq C(e_L) + \mathscr{L}_H(e_L, F_H)$. Hence,

$$\begin{aligned}&[C(e_H)C(e_L)] - [\mathscr{L}_H(e_H, F_H) - \mathscr{L}_L(e_L, F_L)]\\ &\quad = [C(e_H) + \mathscr{L}_H(e_H, F_H)] - [C(e_L) + \mathscr{L}_L(e_L, F_L)]\\ &\quad \leq [\mathscr{L}_H(e_L, F_H) - \mathscr{L}_L(e_L, F_L)].\end{aligned} \tag{B5}$$

From Equation (B3), and the equilibrium identities $\Pi_H = (1 + \pi^*)/2$ and $\Pi_L = \pi^*/2$, $\varepsilon_J = e_J$ we obtain:

$$\begin{aligned}(1 - \pi^*)\frac{\partial \mathscr{L}_H}{\partial \pi^*} &= -(1 - \pi^*)\left[\frac{\alpha_H(1 - eH)}{2}\right]p_H(1/2)\\ &= -\alpha_H(1 - e_H)(1 - \Pi_H)p_H(1/2) = -\mathscr{L}_H(e_H, F_H)\\ \pi^*\frac{\partial \mathscr{L}_L}{\partial \pi^*} &= -(\pi^*)\left[\frac{\alpha_L(1 - e_L)}{2}\right]p_L(1/2)\\ &= -\left[\frac{\pi^*/2}{1 - \Pi_L}\right]\alpha_L(1 - e_L)(1 - \Pi_L)p_L(1/2)\\ &= -\left[\frac{\Pi_L}{1 - \pi_L}\right]\mathscr{L}_L(e_L, F_L).\end{aligned} \tag{B6}$$

Combining (B5) and (B6),

$$\begin{aligned}&[C(e_H) - C(e_L)] - [\mathcal{L}(e_H, F_H) - \mathcal{L}(e_L, F_L)] + \left[(1-\pi^*)\frac{\partial \mathcal{L}_H}{\partial \pi^*} + \pi^* \frac{\partial \mathcal{L}_L}{\partial \pi^*}\right] \\ &\quad \leq -\mathcal{L}_H(e_L, F_H) - \mathcal{L}_L(e_L, F_L) + \mathcal{L}_H(e_H, F_H) + \left[\frac{\Pi_L}{1-\Pi_L}\right]\mathcal{L}_L(e_L, F_L)\end{aligned} \qquad (B7)$$

Next, from equation (11) it follows that:

$$\begin{aligned}\mathcal{L}_H(e_L, F_H) &= \alpha_H(1-\Pi_H)(1-e_L)p_H(1/2) = (1-e_L)C'(e_H) \\ &\geq (1-e_L)C'(e_L) = \alpha_L(1-\Pi_L)(1-e_L)p_L(1/2) = \mathcal{L}_L(e_L, F_L).\end{aligned}$$

Using this relation, we can replace $\mathcal{L}_L(e_L, F_L)$, with $\mathcal{L}_H(e_L, F_H)$ in inequality (B7); it then follows that the right-hand-side of (B4) is positive provided that:

$$0 \leq [F_H - F_L] - [\mathcal{L}_H(eH, F_H) - \mathcal{L}_L(e_L, F_L)] - \left[\frac{1}{1-\Pi_L}\right]\mathcal{L}_H(e_L, F_H). \qquad (B8)$$

The type π^* is indifferent across auditors and in equilibrium we must have beliefs agreeing with strategies, that is, $\Pi_H = (1+\pi^*)/2, \Pi_L = \pi^*/2$, $e_H = \varepsilon_H$ and $e_L = \varepsilon_L$. Substituting into the indifference equation (9) yields:

$$\left[\frac{e_H^2}{d_H} - \frac{e_L^2}{d_L}\right]\left[\pi^* - \frac{1}{2}\right]x + \left[\frac{e_H}{d_H} - \frac{e_L}{d_L}\right]\frac{x}{2} + \left[\frac{(1-e_H)\Pi_H}{d_H} - \frac{(1-e_L)\Pi_L}{d_L}\right]x = [F_H - F_L].$$

By hypothesis, $\pi^* \geq ½$ and the inequality $e_H^2/d_H > e_L^2/d_L$ holds in equilibrium; so the first term on the left-hand side of the equation above is positive. Rearranging the remaining terms, we obtain:

$$\left[\frac{\Pi_H}{d_H} - \left(\frac{e_H}{d_H}\right)\left(\Pi_H - \frac{1}{2}\right)\right]x - \left[\frac{\Pi_L}{d_L} - \left(\frac{e_L}{d_L}\right)\left(\Pi_L - \frac{1}{2}\right)\right]x \leq F_H - F_L$$

In addition, based on the equilibrium identity $d_J = 1 - \alpha_J(1-\Pi_J)(1-e_J)$ for $J = H$ or L, we have:

$$\begin{aligned}\mathcal{L}(e_J, F_J) &= \alpha_J(1-e_J)(1-\Pi_J)p_J(1/2) \\ &= [1-d_J]\left[\frac{\Pi_J}{d_J} - \left(\frac{e_J}{d_J}\right)\left(\Pi_J - \frac{1}{2}\right)\right]x.\end{aligned} \qquad (B9)$$

From the equation above, setting (i) $\Pi_H - \Pi_L = 1/2$ and (ii) replacing e_L in the last term by the larger value e_H, we obtain:

$$[F_H - F_L] - [\mathcal{L}_H(e_H, F_H) - \mathcal{L}_L(e_L, F_L)] \geq \frac{1}{2} - \frac{1}{2}e_H. \qquad (B10)$$

Substituting from (B9) and (B10) into (B8) we need to show that:

$$\begin{aligned}&[F_H - F_L] - [\mathcal{L}_H(e_H, F_H) - \mathcal{L}_L(e_L, F_L)] - \left[\frac{1}{1-\Pi_L}\right]\mathcal{L}_H(e_L, F_H) \\ &\geq \left[\frac{1-e_H}{2}\right] - \left[\frac{1}{1-\Pi_L}\right]\mathcal{L}_H(e_L, F_H) \geq 0.\end{aligned}$$

From the fact that Π_H - Π_L = ½ , it follows that $(1-\Pi_H)/(1-\Pi_L) = (1-\Pi_H)/(3/2-\Pi_H) \leq 1/3$ for values of $\Pi_H \geq 3/4$ (i.e, for $B^* \geq$ ½). Therefore:

$$[1-\Pi_L]^{-1}\mathscr{L}_H(e_L, F_H) = \alpha_H(1-\Pi_H(1-\Pi_L)^{-1}(1-e_L)p_H(1/2) \leq (1/3)p_H(1/2).$$

Differentiation with respect to Π_J shows that (Π_J/d_J) is increasing in Π_J. At the extreme value $\Pi_H = 1$, it is directly verified that $p_H(1/2) = 1-(e_H/2)$. Therefore (B8) reduces to showing that:

$$\left[\frac{1-e_H}{2}\right] - \frac{1}{3}\left[1-\frac{e_H}{2}\right] \geq 0 \Rightarrow \frac{1}{6} - \frac{2}{3}e_H \geq 0.$$

This last inequality follows immediately from the fact $e_H \leq$ ½. Thus, we have established (B8) and thereby (B5). However, (B5) is a contradiction of (B1), that is, we have shown that for $\pi^* >$ ½ (and by implication, $e_H \leq$ ½), we have:

$$0 \leq [F_H - F_L] - \left[\mathscr{L}_H(e_L, F_H) - \mathscr{L}_H(e_L, F_H) + \left(\frac{1}{1-\Pi_L}\right)\mathscr{L}_H(e_L, F_L)\right]$$
$$\leq (1-2\pi^*)\left[\frac{\partial \pi^*}{\partial F_H}\right]^{-1} < 0$$

It follows that $\pi^* \leq$ ½.

Appendix B.4. Proof of Proposition 6

The proof is similar to that of Proposition 5 but uses only the first two equations of (8), that is, profit maximization with regard to fees takes place but the level of inputs in this setting is assumed fixed at some level $e_H = e_L = e^* > 0$. It follows that the rational beliefs are then $\varepsilon_H = \varepsilon_L = e^*$. We also notice that the equality of e_H and e_L automatically implies that the only possible equilibria are ones where the auditor with "big pockets" gets the high end pool (see Proposition 3), that is, $B = H$, $S = L$ and $\Pi_B > \Pi_S$. In addition, it must be the case that $(e^*)^2/d_H > (e^*)^2/d_L$ implying that $1/d_H > 1/d_L$. Further all the equations used in the proof of Proposition 5 with the exception of (B5) and (B7) continue to hold with $e_H = e_L = e^*$ as they do not depend maximization with respect to e. We analyze these equations and show that $\pi^* > 1/3$ can never obtain in equilibrium.

Substituting from (B3) into (B1) and setting $e_H = e_L = e^*$ in yields:

$$[F_H - F_L] - [\mathscr{L}_H(e^*, F_H) - \mathscr{L}_L(e^*, F_L)] - \left[\mathscr{L}_H(e^*, F_H) + \left(\frac{\Pi_L}{1-\Pi_L}\right)\mathscr{L}_L(e^*, F_L)\right]$$
$$= (1-2\pi^*)\left[\frac{\partial \pi^*}{\partial F_H}\right]^{-1}. \tag{B11}$$

Setting $e_H = e_L = e^*$, the indifference equation (9) becomes:

$$[F_H - F_L] = \left[\frac{(e^*)^2}{d_H} - \frac{(e^*)^2}{d_L}\right]\pi^* x + e^*(1-e^*)\left[\frac{1}{d_H} - \frac{1}{d_L}\right]\left[\frac{1}{2}\right]x + (1-e^*)\left[\frac{\Pi_H}{d_H} - \frac{\Pi_L}{d_L}\right]x$$
$$\geq \left[\frac{\partial\pi^*}{\partial F_H}\right]^{-1}\pi^* + (1-e^*)\left[\frac{\Pi_H}{d_H} - \frac{\Pi_L}{d_L}\right]x.$$

Substituting back into (B11) yields:

$$(1-e^*)\left[\frac{\Pi_H}{d_H} - \frac{\Pi_L}{d_L}\right] - \left[2\mathscr{L}_H(e^*, F_H) + \left(\frac{1-2\Pi_L}{1-\Pi_L}\right)\mathscr{L}_L(e^*, F_L)\right]$$
$$\leq (1-3\pi^*)\left[\frac{\partial\pi^*}{\partial F_H}\right]^{-1}. \qquad \text{(B12)}$$

However $\Pi_L \leq 1/2$ and $\Pi_H - \Pi_L = 1/2$. It follows that:

(i) $(1-2\Pi_L)\exists 0$;

(ii) and recalling that $1/d_H \geq 1/d_L(\Pi_H/d_H) - (\Pi_L/d_L) \geq (\Pi_H - \Pi_L)/d_H = 1/2d_H$;

(iii) for $\pi^* \geq 1/3$, we have $\Pi_H \geq 2/3$ and $(1-\Pi_H)\Pi_H = 1/4(1-(\pi^*)^2)2/9$; it follows that $\mathscr{L}_H(e^*, F_H) = \alpha_H(1-\Pi_H)(1-e^*)p_H(1/2) \leq (1-\Pi_H)(1-e^*)p_H(\Pi_H)$ because $\alpha_H \leq 1$ and $\Pi_H \geq ½$. Therefore, $\mathscr{L}_H(e^*, F_H) \leq (1-e^*)(1-\Pi_H)(\Pi_H)/d_H = (1-e^*)2/(9d_H)$.

Using (i), (ii) and (iii) above as well as the fact that $\mathscr{L}_L \exists 0$, we infer from (B12) that:

$$0 < (1-e^*)\left[\frac{1}{2d_H}\right] - (1-e^*)\left[\frac{4}{9d_H}\right] \leq (1-3\pi^*)\left[\frac{\partial\pi^*}{\partial F_H}\right]^{-1}. \qquad \text{(B13)}$$

The right-hand side of (B13) is negative for $\pi^* \geq 1/3$ but the left-hand side is strictly positive (recall that $e^* < 1$). Therefore, we have arrived at a contradiction showing that $\pi^* < 1/3$, that is, auditor $B = H$ takes 2/3 of the market under the assumption that standard setters such a high minimum level of accuracy that it is suboptimal for either auditor to rise above this threshold.

Appendix C: An Example

The parameters are chosen for numerical simplicity of the market share. The choices are:

(1) $\alpha_B = 1$; (2) $\alpha_S = 0$; (3) $k = 0.315$, that is, $C(e) = 0.315\ e^2\ x$.

Because $\alpha_S = 0$ and $\alpha_B > 0$, the last two equations in (8) implies that $e_S = 0$ and $e_B > 0$; consequently, by Proposition 3, $B = H$, $S = L$. It also follows from the parameter choice that:

$$C = (e) = 0.63ex > (\alpha_B/2)ex = 0.5ex$$

and the condition of Proposition 5 is satisfied—therefore, $B = H$ will get the larger market share.

The first-order conditions reduce in this example to:

$$F_B = \frac{e_B^2}{d_B}(1 - \pi^*)x + (0.315 \times e_B^2)x + (1 - \pi^*)(1 - e_B)p_H(1/2),$$
$$F_S = \frac{e_B^2}{d_B}\pi^* x,$$
$$C'(e_B) = (0.63 \times e_B)x = (1 - \pi^*)(1 - e_B)p_H(1/2),$$
$$C'(e_S) = (0.63 \times e_S)x = 0.$$

By direct substitution into the equations above, we verify (within the limits of rounding error) that the equilibrium solution is given by:

$F_B = 0.59\ x$, $e_B = 0.449$; $F_S = 0.06\ x$, $e_S = 0$; $\Pi_B = 3/5$; $\Pi_S = 1/10$; $\pi^* = 1/5$; with associated values: (i) $d_B = 1 - (1 - e_B)(2/5) = 0.79$; (ii) $d_S = 1$; (iii) $p_B(1/2) = 0.71x$ and (iv) $p_S(1/2) = x/10$. The exact solutions are that $e_B = e_H$ solves the quadratic $48\ t^2 - 16t - 3 = 0$ and that $2k = [(1 - e_B)(6 - e_B)(15e_B{}^2 + 10\ e_B)^{-1}]$.

An interesting point about the example is that even if the small auditor has no wealth, the market still gets divided up in equilibrium. The reason for this nontrivial division is the fact that client-firms of low quality are prepared to forgo the benefits of higher insurance in exchange for lower fees. This is best understood by considering the marginal firm π^*. By switching to the high auditor, this marginal firm obtains a higher trading price but the increase in price is exactly offset by the increase in fees.

We conclude this example with a specific illustration of auditor choice by a firm of a given type. In the example above, all firms with cash flows in the range $[0.2x, x]$ choose auditor H who provides audit quality e_B. Consider the implications for a firm whose expected cash flows are $0.4x$, which, under the equilibrium assumption made here, will select auditor $H = B$. Prior to the audit, it is known that the firm's expected cash flow lies in the range $[0.2x, x]$. Given that the auditor sets quality $e_B = 0.452$, the correct value, $0.4x$, is identified and reported by the auditor with probability 0.452. However, with probability $(1 - 0.452) = 0.548$, the firms cash flows are incorrectly assessed by the auditor who proceeds to report this incorrect assessment. In other words, auditors always report their actual assessment of the expected cash flows but with probability $(1-e_B)$, their assessments are in error. Our simplifying assumption is that in the case of audit error, all assessments in the range $[0.2x, x]$ are equally likely. Given this simplifying assumption, conditional on audit error, the joint distribution of firm-types and reports is uniform over the region $[0.2x, x]$ H $[0.2x, x]$. An application of Bayes' rule shows that conditional on an audit error, the expected type of the firm is always the average value, $0.6x$ for every report. Thus, a report of $0.4x$ issued by auditor H would in fact lead to investors to expected cash flow of $(0.452)(0.4x) + (0.548)(0.6x)$.

A more general framework would allow for some arbitrary density $f(\theta|\varphi,e)$ over erroneous reports given true type φ, or equivalently, a joint density $f(\varphi,\theta|e)$. For instance, a firm of true type φ would expect a report $\theta = \varphi$ with probability e and a report θ different from φ with probability $(1-e)f(\theta|\varphi,e)$. Because φ is distributed uniformly, the posterior density of types given reports: $g(\phi|\theta,e) = \text{J} f(\phi,\theta|e)\ d\theta$. The expected cash flows on observing a report of θ would then be: $e\theta\text{x} + (1-e)\ E_g[\phi|\theta]\text{x}$. Such a generalization would have no qualitative impact for our results.

Acknowledgments

The authors would like to thank Jeffrey Callen, Don Campbell, Ron Dye, Mark Finn, Joe Kerstein, Uri Ronnen, Sri Sridharan, Itzhak Venezia and Dror Zuckerman, Stefan Reichelstein and two anonymous referees of this journal. Special thanks are due to Stan Baiman whose comments were invaluable. Bar-Yosef would like to acknowledge financial support from the Krueger Center and the Center for Accounting Studies at the Hebrew University. We also wish to thank seminar participants at the EAA meetings in Bergen, London School of Business, New York University, Northwestern University, and University of Toronto.

Notes

1. For example, Palmrose in her studies of audit fees structure (1986, 1989), reports that out of 361 sample firms spanning the years (1980–1981), 298 firms (83%) were audited by "Big Six" auditors.
2. Our analysis assumes that auditors are initially endowed with unequal wealth levels. We note that the the initial asymmetry in auditors wealth levels will be sustained even when one assumes capital market conditions where borrowing is unrestricted but the less-capitalized auditor faces a higher cost of capital. Under such conditions, this auditor's borrowing cost could exceed the benefits from becoming well capitalized, making it optimal for the auditor to remain small.
3. As will be emphasized later, we abstract away from cost issues and assume that all firms may be audited to the same degree of accuracy at the same cost. This is a fairly standard assumption that does not have any real significance for our analysis.
4. There are two points to clarify here. First, we do not treat GAAP and GAAS separately when we talk of standards, but rather, the combined effect on the accuracy of audit reports. Second, our formal analysis is conducted inside a static context where we assume the existence of two types of auditors with differential capacity to pay damages. As a consequence we do not analyze dynamic issues like entry of new audit firms. In other words, our results provide some insights into factors that affect changes in market shares over time but we do not provide a complete analysis of the dynamics of market share evolution.
5. We use the term "boutique" to characterize a service firm, which has small capitalization (relative to its competition) but offers high quality services at high prices and attracts the best clients (who constitute a small share of the market). Boutiques are common in other industries such as investment banking where clients are provided with services that directly affect their financial outcomes. In contrast, the audit industry repackages information supplied to them by client-firms and provides a service to third-party investors. To summarize, unlike other service industries, the primary consumers of the audited financial statement, the investors, have no direct interaction with the auditor and have to rely on *perceptions* as to whether a quality audit was supplied.
6. One key economic feature that acts against the survival of "boutique" auditors is that audit quality cannot be credibly communicated prior to trade. In contrast, Investment bankers often have personal relationships with the providers of capital and explain their valuation methodology in great detail when trying to raise capital. Consequently, they are able to communicate the quality of their work in a credible fashion prior to trade and our analysis precluding "boutiques" does not apply to the investment banking industry.

7. We assume that the owner sells the entire firm (as in Dye, 1993, 1995), as we primarily focus on the role of the auditor in the valuation process. A partial sale of the firm may signal the value of the client-firm to the market (see Leland and Pyle, 1977) considerably complicates the informational role of audits (see Datar et al., 1991). However, such signaling issues do not have much of a bearing on the analysis of competition for clients by auditors which is the main issue addressed in this paper. However signaling quality through auditor choice is another possibility (see for example Bar-Yosef and Livnat, 1985), which is incorporated into our analysis.
8. There has been considerable discussion, particularly in the empirical literature, regarding the variation of services and fees across audit firms (e.g., Simunic, 1980; Palmrose, 1986). Associated theoretical issues may be found in Antle and Demski (1991).
9. Ronen (1994) also considers the investment in quality adopted by profit-maximizing auditors. However, his model, unlike ours, assumes that all auditors are identical and that adverse selection is absent from the market. Ronen analyzes optimal quality levels chosen by identical auditors whereas we analyze the effects of unequal wealth levels across auditors.
10. See footnote 3. Notice that despite identical cost curves, the equilibrium costs of audits do differ because of differential choices regarding the quality of audits.
11. Alternatively, this amount may be considered to be the accountant's co-payment in a settlement that is primarily made by the insurer. More generally, the cost to the auditor, given a suit and settlement, should include the expected value of future premium increases. Note that the proportion recovered, $\alpha(e)$, is dependent on the auditor type, J, and the care expended on the audit, e. This functional form allows both for negligence rules ($\alpha(e) = 0$ for e $>$ e^* where e^* is some prespecified level of due care), strict liability rules ($\alpha(e)$ independent of e), and proportionate liability ($\alpha(e) = \gamma(e)\ H\Delta(e)$ where γ is the probability of being found guilty and Δ is the proportion of the damages borne by the auditor).
12. We are assuming here that the prior beliefs of the auditors agrees with those of investors so the superscript i used to indicate investors' beliefs applies to auditors also. This is discussed further on page 11.
13. For example, if the market believes that auditors are chosen at random, then $\Pi_B = \Pi_S = 1/2$. In general, the whole distribution of beliefs regarding the types selecting each auditor may be relevant. However, only the average of these beliefs is needed for determining the market prices in our context.
14. This characterization affects both bankruptcy risk and firm value through a single parameter. A more general formulation where both cash flows and firm type are unknown is conceptually quite similar, but results are technically very different.
15. We assume that ϕ, the information obtained by the client-firm, is imperfect (with regard to the true type ω), both for descriptive appeal and for the following technical reason. A rigorous derivation of the equilibrium requires that every report θ occurs with positive probability regardless of how the market segments across auditors. Making the firm's private information imperfect and assuming that the probability density of the true type contingent on the signal is strictly positive, i.e., $f(\omega\mid\phi) > 0$, allows us to ensure that every possible audit report occurs with some positive probability for every choice of F and e. That is, whatever choice of F and e and the consequent segmentation of the market, there is no out-of-equilibrium report θ and no technical problems regarding out-of equilibrium beliefs (see also footnote 23).
16. Consistent with auditor independence, notice that auditors report their actual observation without strategic distortions. This provides us with a considerable simplification and enables us to obtain a closed-form solution for the market price in terms of the report. Given this assumption, the nature of the error in the audit process is not critical, that is, the assumption that the report is uninformative given the error can be replaced with more general distributions of true firm-types given audit error. However, if auditors collude with client-firms and strategically distort the audit report, the consequences for our paper are unpredictable.
17. We assume that projects are sufficiently profitable so the owners can borrow the amount needed to finance the project and pay auditors. Any excess cash left over after investment in the project and payment of audit fees is paid out as dividends to current owners. However, if the capital is insufficient (to finance the investment and to pay the audit fee), the firm borrows the additional required cash and repays it from the proceeds of the sale to the new owners. Thus, the new shareholders simply obtain the cash flows, x, from the uncertain project/asset whereas the old owners receive the net of the sale

price, p, and any debt they might have assumed before starting on the project. In addition, this equation uses the fact that beliefs regarding the level of audit accuracy, ε, are independent of the report θ. Such independence follows from our assumption that conditional on any erroneous report θ, the posterior distribution of firm types is uniformly distributed over the client-firm types choosing auditor J.

18. More generally, the recovery fraction will also depend on the level of inputs, e. In contrast, our formulation assumes that if an audit error is established, the level of care does not affect the damages. The argument made is that the courts respond to an auditor error by saying: "it does not matter how hard you say you worked, you did not work hard enough to avoid the mistake."
19. We are assuming here that the probability ε_J, of correctly identifying ω, is independent of ω, that is, for a given level of audit care, all firms have equal probabilities of being misidentified.
20. A technical feature underlying this proof is the assumption that firms are atomistic ensuring that each individual firm's choice of auditor leaves the average failure risk (second term in equation (6) unchanged. Because the difference in fees (last term in equation (6) is also unaffected by the firm's auditor selection, auditor choice is determined by the first term in equation (6) which is a function of the beliefs regarding audit quality.
21. We will allocate the actual indifferent client-firm type, π^*, to auditor H as a matter of convention. Our results are unaffected by whether π^* chooses H or L or randomizes across the two auditor types.
22. The justification for this substitution is provided in Appendix A where we formally derive the equilibrium.
23. The formal structure has been chosen carefully to ensure that neither the report θ nor the audit fee F provide information about the input choice e. Specifically, the fact that $f(\omega|\phi) > 0$ for every ω, ϕ ensures that every report θ has positive probability irrespective of the choices of e and F. We also assume that client-firms cannot infer the optimal level of inputs, e, that correspond with a particular level of F, that is, observing F does not convey information regarding e. This condition will hold, for example, if client-firms do not know the auditor's cost function.
24. Specifically, the condition we impose on the cost function in Proposition 5 ensures that the accuracy level of the auditor attracting high-type firms drops below 1/2 as soon as the minimum client-firm type π^* becomes greater than 1/2 (see Appendix B.3). There are many alternative conditions linking the cost function and litigation cost that will also ensure the result in Proposition 5. However, all of them will necessitate that the quality level drops "sufficiently" fast as π^* increases. The specific one that we have selected has the merit that it is easy to state and leads relatively quickly to a bound on e as a function of π^*.
25. This penalty structure is very similar to the one used in Dye (1993).
26. We abstract away from strategic behavior and manipulation of the type discussed in Dye (2002).

References

Antle, R. (1982). "The Auditor as an Economic Agent." *Journal of Accounting Research* 20:2 , 503–527.

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

Arthur Anderson & Co., Coopers & Lybrand, Deloitte & Touche, Ernst & Young, KPMG Peat Marwick, and Price Waterhouse (1992) "The Liability Crisis in the United States: Impact on the Accounting Profession."

Bar-Yosef, S. and J. Livnat. (1985). "Auditor Selection: An Incentive-Signaling Approach." *Accounting and Business Research* 14:56, 301–309.

Baiman, S., J. Evans and J. Noel. (1987). "Optimal Contracts with a Utility-Maximizing Auditor." *Journal of Accounting Research* 25, 217–224.

Beatty, R. P. (1989). "Auditor Reputation and the Pricing of Initial Public Offerings." *The Accounting Review* 64:4, 693–709.

Cho, I. and D. M. Kreps. (1987). "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 180–221.

Datar, S. M., G. A. Feltham and J. S. Hughes. (1991). "The Role of Audits and Audit Quality in Valuing new Issues," *Journal of Accounting and Economics* 14:1, 3–49.

DeAngelo, L. (1981). "Auditor Size and Audit Quality." *Journal of Accounting and Economics* 4, 183–199.

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

Dye, R. A. (1995). "Incorporation and the Audit Market." *Journal of Accounting and Economics* 18, 75–114.

Dye, R. A. (2002). "Classifications Manipulation and Nash Accounting Standards." *Journal of Accounting Research* 40:4.

Leland, H. E. and D. H. Pyle. (1977). "Information Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance* 32, 371–387.

Mansi, S. A., W. F. Maxwell and D. D. Miller. (2003). "Does Auditors Quality and Tenure Matter to Investors? Evidence from the Bond Market." SSRN_ ID384594.

Menon, K. and Williams. (1994). "The Insurance Hypothesis and Market Prices." *The Accounting Review* 327–342.

Palmrose, Z. (1986). "Audit Fees and Auditor Size: Further Evidence." *Journal of Accounting Research* 24, 97–110.

Palmrose, Z. (1989). "The Relation of Audit Contract Type to Audit Fees and Hours." *The Accounting Review* 64, 488–499.

Ratner, D. L. (1988). *Securities Regulation*. West Publishing Company.

Ronen, U. (1994). "The Effects of a Mandatory Auditing Policy on the Quality of Auditing." Working Paper, Rutgers University.

Simunic, D. A. (1980). "The Pricing of Audit Services: Theory and Evidence." *Journal of Accounting Research* 18, 161–190.

Titman, S. and B. Trueman (1986). "Information Quality and The Valuation of New Issues." *Journal of Accounting and Economics* 8, 159–172.